

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2005, 19:44:24 ; Search time 16.3256 Seconds
(without alignments)
59.443 Million cell updates/sec

Title: US-10-812-238B-23

Perfect score: 73

Sequence: 1 NTRCRGDSKVQE 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	73	100.0	233	4	US-09-360-376-55
2	44	60.3	705	4	US-08-311-731A-4
3	40	54.8	282	4	US-09-360-376-54
4	40	54.8	424	4	US-09-489-039A-12030
5	40	54.8	501	4	US-09-270-767-41177
6	40	54.8	501	4	US-09-270-767-56393
7	39	53.4	283	3	US-08-992-035A-3
8	39	53.4	283	4	US-09-360-376-53
9	38	52.1	120	4	US-09-513-999C-6407
10	38	52.1	136	4	US-09-690-454-131
11	37	50.7	201	3	US-09-015-734-12
12	37	50.7	201	4	US-09-015-311-12
13	37	50.7	236	3	US-09-015-734-7
14	37	50.7	236	4	US-09-015-311-7
15	37	50.7	255	3	US-09-015-734-2
16	37	50.7	255	4	US-09-015-311-2
17	37	50.7	272	4	US-09-489-039A-13748
18	37	50.7	275	4	US-09-489-039A-11789
19	37	50.7	488	4	US-09-489-039A-13164
20	36	49.3	280	4	US-09-601-478-5
21	36	49.3	280	4	US-09-601-478-8
22	36	49.3	280	4	US-09-442-013-13
23	36	49.3	280	4	US-09-513-365A-1
24	36	49.3	280	4	US-09-513-365A-6
25	36	49.3	316	4	US-09-328-352-8163
26	36	49.3	416	4	US-09-949-016-8237
27	36	49.3	498	4	US-09-786-240-1
					Sequence 1, Appli
					Sequence 55, Appli
					Sequence 4, Appli
					Sequence 54, Appli
					Sequence 12030, A
					Sequence 41177, A
					Sequence 56393, A
					Sequence 3, Appli
					Sequence 53, Appli
					Sequence 6407, Ap
					Sequence 131, App
					Sequence 12, Appl
					Sequence 7, Appli
					Sequence 2, Appli
					Sequence 8, Appli
					Sequence 13748, A
					Sequence 11789, A
					Sequence 13164, A
					Sequence 5, Appli
					Sequence 8, Appli
					Sequence 13, Appl
					Sequence 1, Appli
					Sequence 8163, Ap
					Sequence 8237, Ap
					Sequence 1, Appli
					Sequence 1290, Ap
					Sequence 12, Appl
					Sequence 12, Appl
					Sequence 22600, A
					Sequence 78, Appl
					Sequence 6121, Ap
					Sequence 6, Appli
					Sequence 7261, Ap
					Sequence 1, Appli
					Sequence 92, Appl
					Sequence 95, Appl
					Sequence 92, Appl
					Sequence 95, Appl
					Sequence 12, Appl
					Sequence 8, Appli
					Sequence 12, Appl
					Sequence 10, Appl
					Sequence 17110, A
					Sequence 4, Appli
					Sequence 3, Appli
					Sequence 3, Appli
					Sequence 2, Appli
					Sequence 4, Appli
					Sequence 4, Appli
					Sequence 2, Appli
					Sequence 6624, Ap
					Sequence 10092, A
					Sequence 23746, A
					Sequence 107, App
					Sequence 107, App
					Sequence 11044, A
					Sequence 6515, Ap
					Sequence 10331, A
					Sequence 25139, A
					Sequence 1291, Ap
					Sequence 86, Appl
					Sequence 18617, A
					Sequence 22192, A
					Sequence 59, Appl
					Sequence 59, Appl
					Sequence 10, Appl
					Sequence 59, Appl
					Sequence 59, Appl
					Sequence 8206, Ap
					Sequence 59, Appl
					Sequence 59, Appl
					Sequence 15729, A
					Sequence 4948, Ap
					Sequence 9, Appli
					Sequence 5053, Ap
					Sequence 7, Appli
					Sequence 30671, A
					Sequence 21699, A
					Sequence 86, Appl
					Sequence 89, Appl
					Sequence 86, Appl
					Sequence 89, Appl
					Sequence 18, Appl
					Sequence 19, Appl
					Sequence 14, Appl
					Sequence 11, Appl
					Sequence 34, Appl
					Sequence 104, Appl

101	34	46.6	757	4	US-10-065-133A-104	Sequence 104, App	174	33	45.2	35	4	US-09-515-965A-140	Sequence 140, App
102	34	46.6	971	3	US-09-107-149-19	Sequence 19, App	175	33	45.2	35	4	US-09-515-965A-141	Sequence 141, App
103	34	46.6	1036	3	US-08-968-752B-4	Sequence 4, Appl	176	33	45.2	35	4	US-09-515-965A-142	Sequence 142, App
104	34	46.6	1036	3	US-09-536-224-4	Sequence 4, Appl	177	33	45.2	35	4	US-09-515-965A-143	Sequence 143, App
105	34	46.6	1074	2	US-08-768-147B-2	Sequence 2, Appl	178	33	45.2	35	4	US-09-515-965A-144	Sequence 144, App
106	34	46.6	1074	3	US-08-968-752B-2	Sequence 2, Appl	179	33	45.2	35	4	US-09-515-965A-145	Sequence 145, App
107	34	46.6	1074	3	US-09-107-149-3	Sequence 3, Appl	180	33	45.2	35	4	US-09-515-965A-146	Sequence 146, App
108	34	46.6	1074	3	US-09-536-224-2	Sequence 2, Appl	181	33	45.2	35	4	US-09-515-965A-147	Sequence 147, App
109	34	46.6	1086	4	US-09-949-016-11159	Sequence 11159, A	182	33	45.2	35	4	US-09-515-965A-148	Sequence 148, App
110	33	45.2	28	3	US-09-082-279B-18	Sequence 18, Appl	183	33	45.2	35	4	US-09-515-965A-149	Sequence 149, App
111	33	45.2	28	3	US-09-082-279B-19	Sequence 19, Appl	184	33	45.2	35	4	US-09-515-965A-150	Sequence 150, App
112	33	45.2	28	3	US-08-474-349A-234	Sequence 224, App	185	33	45.2	35	4	US-09-515-965A-151	Sequence 151, App
113	33	45.2	28	3	US-08-474-349A-235	Sequence 225, App	186	33	45.2	35	4	US-09-350-641C-22	Sequence 22, Appl
114	33	45.2	28	3	US-09-315-304B-18	Sequence 18, Appl	187	33	45.2	35	4	US-09-350-641C-138	Sequence 138, App
115	33	45.2	28	3	US-09-315-304B-19	Sequence 19, Appl	188	33	45.2	35	4	US-09-350-641C-139	Sequence 139, App
116	33	45.2	28	4	US-09-834-784-18	Sequence 18, Appl	189	33	45.2	35	4	US-09-350-641C-140	Sequence 140, App
117	33	45.2	28	4	US-09-834-784-19	Sequence 19, Appl	190	33	45.2	35	4	US-09-350-641C-141	Sequence 141, App
118	33	45.2	28	4	US-09-515-965A-18	Sequence 18, Appl	191	33	45.2	35	4	US-09-350-641C-142	Sequence 142, App
119	33	45.2	28	4	US-09-515-965A-19	Sequence 19, Appl	192	33	45.2	35	4	US-09-350-641C-143	Sequence 143, App
120	33	45.2	28	4	US-09-350-641C-18	Sequence 18, Appl	193	33	45.2	35	4	US-09-350-641C-144	Sequence 144, App
121	33	45.2	28	4	US-09-350-641C-19	Sequence 19, Appl	194	33	45.2	35	4	US-09-350-641C-145	Sequence 145, App
122	33	45.2	28	4	US-09-350-841A-18	Sequence 18, Appl	195	33	45.2	35	4	US-09-350-641C-146	Sequence 146, App
123	33	45.2	28	4	US-09-350-841A-19	Sequence 19, Appl	196	33	45.2	35	4	US-09-350-641C-147	Sequence 147, App
124	33	45.2	35	3	US-08-485-264A-210	Sequence 210, App	197	33	45.2	35	4	US-09-350-641C-148	Sequence 148, App
125	33	45.2	35	3	US-09-082-279B-22	Sequence 22, Appl	198	33	45.2	35	4	US-09-350-641C-149	Sequence 149, App
126	33	45.2	35	3	US-09-082-279B-138	Sequence 138, App	199	33	45.2	35	4	US-09-350-641C-150	Sequence 150, App
127	33	45.2	35	3	US-09-082-279B-139	Sequence 139, App	200	33	45.2	35	4	US-09-350-641C-151	Sequence 151, App
128	33	45.2	35	3	US-09-082-279B-140	Sequence 140, App	201	33	45.2	35	4	US-09-350-841A-22	Sequence 22, Appl
129	33	45.2	35	3	US-09-082-279B-141	Sequence 141, App	202	33	45.2	35	4	US-09-350-841A-138	Sequence 138, App
130	33	45.2	35	3	US-09-082-279B-142	Sequence 142, App	203	33	45.2	35	4	US-09-350-841A-139	Sequence 139, App
131	33	45.2	35	3	US-09-082-279B-143	Sequence 143, App	204	33	45.2	35	4	US-09-350-841A-140	Sequence 140, App
132	33	45.2	35	3	US-09-082-279B-144	Sequence 144, App	205	33	45.2	35	4	US-09-350-841A-141	Sequence 141, App
133	33	45.2	35	3	US-09-082-279B-145	Sequence 145, App	206	33	45.2	35	4	US-09-350-841A-142	Sequence 142, App
134	33	45.2	35	3	US-09-082-279B-146	Sequence 146, App	207	33	45.2	35	4	US-09-350-841A-143	Sequence 143, App
135	33	45.2	35	3	US-09-082-279B-147	Sequence 147, App	208	33	45.2	35	4	US-09-350-841A-144	Sequence 144, App
136	33	45.2	35	3	US-09-082-279B-148	Sequence 148, App	209	33	45.2	35	4	US-09-350-841A-145	Sequence 145, App
137	33	45.2	35	3	US-09-082-279B-149	Sequence 149, App	210	33	45.2	35	4	US-09-350-841A-146	Sequence 146, App
138	33	45.2	35	3	US-09-082-279B-150	Sequence 150, App	211	33	45.2	35	4	US-09-350-841A-147	Sequence 147, App
139	33	45.2	35	3	US-09-082-279B-151	Sequence 151, App	212	33	45.2	35	4	US-09-350-841A-148	Sequence 148, App
140	33	45.2	35	3	US-08-474-349A-228	Sequence 228, App	213	33	45.2	35	4	US-09-350-841A-149	Sequence 149, App
141	33	45.2	35	3	US-09-315-304B-22	Sequence 22, Appl	214	33	45.2	35	4	US-09-350-841A-150	Sequence 150, App
142	33	45.2	35	3	US-09-315-304B-138	Sequence 138, App	215	33	45.2	35	4	US-09-350-841A-151	Sequence 151, App
143	33	45.2	35	3	US-09-315-304B-139	Sequence 139, App	216	33	45.2	37	3	US-08-919-597-210	Sequence 210, App
144	33	45.2	35	3	US-09-315-304B-140	Sequence 140, App	217	33	45.2	37	3	US-08-919-597-211	Sequence 211, App
145	33	45.2	35	3	US-09-315-304B-141	Sequence 141, App	218	33	45.2	37	3	US-08-919-597-212	Sequence 212, App
146	33	45.2	35	3	US-09-315-304B-142	Sequence 142, App	219	33	45.2	37	4	US-08-470-896-210	Sequence 210, App
147	33	45.2	35	3	US-09-315-304B-143	Sequence 143, App	220	33	45.2	37	4	US-08-470-896-211	Sequence 211, App
148	33	45.2	35	3	US-09-315-304B-144	Sequence 144, App	221	33	45.2	37	4	US-08-470-896-212	Sequence 212, App
149	33	45.2	35	3	US-09-315-304B-145	Sequence 145, App	222	33	45.2	42	3	US-08-486-099-121	Sequence 121, App
150	33	45.2	35	3	US-09-315-304B-146	Sequence 146, App	223	33	45.2	42	3	US-08-484-223B-121	Sequence 121, App
151	33	45.2	35	3	US-09-315-304B-147	Sequence 147, App	224	33	45.2	42	3	US-08-919-597-121	Sequence 121, App
152	33	45.2	35	3	US-09-315-304B-148	Sequence 148, App	225	33	45.2	42	3	US-08-475-668A-121	Sequence 121, App
153	33	45.2	35	3	US-09-315-304B-149	Sequence 149, App	226	33	45.2	42	3	US-08-485-551A-121	Sequence 121, App
154	33	45.2	35	3	US-09-315-304B-150	Sequence 150, App	227	33	45.2	42	3	US-08-471-913A-121	Sequence 121, App
155	33	45.2	35	3	US-09-315-304B-151	Sequence 151, App	228	33	45.2	42	3	US-08-485-264A-121	Sequence 121, App
156	33	45.2	35	4	US-09-834-784-22	Sequence 22, Appl	229	33	45.2	42	3	US-09-082-279B-17	Sequence 17, Appl
157	33	45.2	35	4	US-09-834-784-138	Sequence 138, App	230	33	45.2	42	3	US-08-474-349A-121	Sequence 121, App
158	33	45.2	35	4	US-09-834-784-139	Sequence 139, App	231	33	45.2	42	3	US-09-315-304B-17	Sequence 17, Appl
159	33	45.2	35	4	US-09-834-784-140	Sequence 140, App	232	33	45.2	42	4	US-08-487-266A-121	Sequence 121, App
160	33	45.2	35	4	US-09-834-784-141	Sequence 141, App	233	33	45.2	42	4	US-08-485-546A-121	Sequence 121, App
161	33	45.2	35	4	US-09-834-784-142	Sequence 142, App	234	33	45.2	42	4	US-09-834-784-17	Sequence 17, Appl
162	33	45.2	35	4	US-09-834-784-143	Sequence 143, App	235	33	45.2	42	4	US-09-515-965A-17	Sequence 17, Appl
163	33	45.2	35	4	US-09-834-784-144	Sequence 144, App	236	33	45.2	42	4	US-09-350-641C-17	Sequence 17, Appl
164	33	45.2	35	4	US-09-834-784-145	Sequence 145, App	237	33	45.2	42	4	US-09-350-841A-17	Sequence 17, Appl
165	33	45.2	35	4	US-09-834-784-146	Sequence 146, App	238	33	45.2	42	4	US-08-487-266A-121	Sequence 121, App
166	33	45.2	35	4	US-09-834-784-147	Sequence 147, App	239	33	45.2	48	3	US-08-486-099-16	Sequence 16, Appl
167	33	45.2	35	4	US-09-834-784-148	Sequence 148, App	240	33	45.2	48	3	US-08-360-107A-16	Sequence 16, Appl
168	33	45.2	35	4	US-09-834-784-149	Sequence 149, App	241	33	45.2	48	3	US-08-484-223B-16	Sequence 16, Appl
169	33	45.2	35	4	US-09-834-784-150	Sequence 150, App	242	33	45.2	48	3	US-08-919-597-16	Sequence 16, Appl
170	33	45.2	35	4	US-09-834-784-151	Sequence 151, App	243	33	45.2	48	3	US-08-475-668A-16	Sequence 16, Appl
171	33	45.2	35	4	US-09-515-965A-22	Sequence 22, Appl	244	33	45.2	48	3	US-08-485-551A-16	Sequence 16, Appl
172	33	45.2	35	4	US-09-515-965A-138	Sequence 138, App	245	33	45.2	48	3	US-08-471-913A-16	Sequence 16, Appl
173	33	45.2	35	4	US-09-515-965A-139	Sequence 139, App	246	33	45.2	48	3	US-08-485-264A-16	Sequence 16, Appl

247	33	45.2	48	3	US-08-474-349A-16	Sequence 16, Appl	320	33	45.2	730	2	US-09-060-692-58	Sequence 58, Appl
248	33	45.2	48	4	US-08-455-208A-33	Sequence 33, Appl	321	33	45.2	730	3	US-08-833-391-58	Sequence 58, Appl
249	33	45.2	48	4	US-08-470-896-16	Sequence 16, Appl	322	33	45.2	730	3	US-09-060-610-58	Sequence 58, Appl
250	33	45.2	48	4	US-08-485-546A-16	Sequence 16, Appl	323	33	45.2	730	5	PCT-US94-10151A-58	Sequence 58, Appl
251	33	45.2	48	4	US-08-487-266A-16	Sequence 16, Appl	324	33	45.2	754	3	US-09-276-400-8	Sequence 8, Appl
252	33	45.2	50	4	US-08-255-208A-16	Sequence 16, Appl	325	33	45.2	754	3	US-09-448-076-8	Sequence 8, Appl
253	33	45.2	102	3	US-09-034-916-5	Sequence 5, Appl	326	33	45.2	754	3	US-09-702-572-8	Sequence 8, Appl
254	33	45.2	103	4	US-09-107-532A-7072	Sequence 7072, Ap	327	33	45.2	800	4	US-09-489-039A-7191	Sequence 7191, Ap
255	33	45.2	115	4	US-09-270-767-40288	Sequence 40288, A	328	33	45.2	825	4	US-09-489-039A-8469	Sequence 8469, Ap
256	33	45.2	115	4	US-09-270-767-55504	Sequence 55504, A	329	33	45.2	850	4	US-09-107-532A-4631	Sequence 4631, Ap
257	33	45.2	130	3	US-08-329-799-37	Sequence 37, Appl	330	33	45.2	1048	4	US-09-949-016-6192	Sequence 6192, Ap
258	33	45.2	158	4	US-09-248-796A-18178	Sequence 18178, A	331	33	45.2	1149	4	US-09-949-016-7401	Sequence 7401, Ap
259	33	45.2	174	4	US-09-107-532A-4715	Sequence 4715, Ap	332	33	45.2	1436	4	US-09-578-063-78	Sequence 78, Appl
260	33	45.2	253	4	US-09-252-991A-32459	Sequence 32459, A	333	33	45.2	3200	2	US-08-477-451-8	Sequence 8, Appl
261	33	45.2	299	4	US-09-252-991A-21857	Sequence 21857, A	334	33	43.8	7	1	US-08-421-702A-17	Sequence 17, Appl
262	33	45.2	308	4	US-09-248-796A-14594	Sequence 14594, A	335	33	43.8	7	1	US-08-303-052A-17	Sequence 17, Appl
263	33	45.2	310	1	US-07-704-288C-6	Sequence 6, Appl	336	32	43.8	7	1	US-08-421-696A-17	Sequence 17, Appl
264	33	45.2	310	1	US-08-379-259-6	Sequence 6, Appl	337	32	43.8	7	1	US-08-421-697A-17	Sequence 17, Appl
265	33	45.2	310	3	US-07-791-931-6	Sequence 6, Appl	338	32	43.8	7	1	US-08-421-698A-17	Sequence 17, Appl
266	33	45.2	313	4	US-09-302-540-11218	Sequence 11218, A	339	32	43.8	7	2	US-08-421-695A-17	Sequence 17, Appl
267	33	45.2	324	1	US-08-047-413-11	Sequence 11, Appl	340	32	43.8	7	5	PCT-US95-04741-17	Sequence 17, Appl
268	33	45.2	324	3	US-08-229-050-11	Sequence 11, Appl	341	32	43.8	9	1	US-08-421-702A-26	Sequence 26, Appl
269	33	45.2	324	3	US-08-801-563-11	Sequence 11, Appl	342	32	43.8	9	1	US-08-421-696A-26	Sequence 26, Appl
270	33	45.2	346	3	US-09-034-916-2	Sequence 2, Appl	343	32	43.8	9	1	US-08-421-697A-26	Sequence 26, Appl
271	33	45.2	347	4	US-09-307-794A-148	Sequence 148, App	344	32	43.8	9	1	US-08-421-698A-26	Sequence 26, Appl
272	33	45.2	347	4	US-09-305-125A-148	Sequence 148, App	345	32	43.8	9	1	US-08-421-695A-26	Sequence 26, Appl
273	33	45.2	347	4	US-09-302-775A-148	Sequence 148, App	346	32	43.8	9	5	PCT-US95-04741-26	Sequence 26, Appl
274	33	45.2	347	4	US-09-306-700-148	Sequence 148, App	347	32	43.8	20	3	US-08-857-076-27	Sequence 27, Appl
275	33	45.2	347	4	US-09-303-603A-148	Sequence 148, App	348	32	43.8	35	3	US-08-857-076-30	Sequence 30, Appl
276	33	45.2	347	4	US-09-304-920A-148	Sequence 148, App	349	32	43.8	43	3	US-08-857-076-28	Sequence 28, Appl
277	33	45.2	347	4	US-09-309-064-148	Sequence 148, App	350	32	43.8	65	4	US-09-248-798A-27954	Sequence 27954, A
278	33	45.2	347	4	US-09-305-381A-148	Sequence 148, App	351	32	43.8	72	4	US-09-328-352-4610	Sequence 4610, Ap
279	33	45.2	347	4	US-09-306-618-148	Sequence 148, App	352	32	43.8	90	4	US-09-270-767-31781	Sequence 31781, A
280	33	45.2	359	4	US-09-372-758A-2	Sequence 2, Appl	353	32	43.8	90	4	US-09-270-767-38050	Sequence 38050, A
281	33	45.2	362	4	US-09-252-991A-24409	Sequence 24409, A	354	32	43.8	90	4	US-09-270-767-46998	Sequence 46998, A
282	33	45.2	369	4	US-09-949-016-10311	Sequence 10311, A	355	32	43.8	90	4	US-09-270-767-53267	Sequence 53267, A
283	33	45.2	383	4	US-09-248-796A-18954	Sequence 18954, A	356	32	43.8	98	4	US-09-513-999C-4239	Sequence 4239, Ap
284	33	45.2	389	4	US-09-270-767-32291	Sequence 32291, A	357	32	43.8	100	4	US-09-621-976-4269	Sequence 4269, Ap
285	33	45.2	389	4	US-09-270-767-47508	Sequence 47508, A	358	32	43.8	100	4	US-09-513-999C-5033	Sequence 5033, Ap
286	33	45.2	469	4	US-09-347-650-16	Sequence 16, Appl	359	32	43.8	107	4	US-09-107-532A-5073	Sequence 5073, Ap
287	33	45.2	503	4	US-09-302-540-11343	Sequence 11343, A	360	32	43.8	118	4	US-08-794-028B-21	Sequence 21, Appl
288	33	45.2	527	2	US-08-659-939-4	Sequence 4, Appl	361	32	43.8	118	4	US-08-794-028B-22	Sequence 22, Appl
289	33	45.2	527	2	US-08-850-041-4	Sequence 4, Appl	362	32	43.8	119	3	US-08-753-642-2	Sequence 2, Appl
290	33	45.2	527	3	US-08-896-500-4	Sequence 4, Appl	363	32	43.8	120	1	US-07-979-630-1	Sequence 1, Appl
291	33	45.2	527	3	US-08-476-397-4	Sequence 4, Appl	364	32	43.8	120	1	US-08-440-049-1	Sequence 1, Appl
292	33	45.2	527	3	US-08-973-720-4	Sequence 4, Appl	365	32	43.8	120	1	US-08-440-049-3	Sequence 3, Appl
293	33	45.2	527	3	US-08-923-558-2	Sequence 2, Appl	366	32	43.8	120	2	US-08-441-513A-1	Sequence 1, Appl
294	33	45.2	527	3	US-09-262-927-4	Sequence 2, Appl	367	32	43.8	120	2	US-08-441-513A-3	Sequence 3, Appl
295	33	45.2	527	4	US-09-486-553-2	Sequence 2, Appl	368	32	43.8	120	3	US-08-970-865-2	Sequence 2, Appl
296	33	45.2	527	4	US-09-570-383-4	Sequence 4, Appl	369	32	43.8	120	3	US-08-970-865-3	Sequence 3, Appl
297	33	45.2	527	4	US-09-395-336-4	Sequence 4, Appl	370	32	43.8	120	3	US-08-581-662-31	Sequence 31, Appl
298	33	45.2	574	6	5223254-2	Patent No. 5223254	371	32	43.8	120	3	US-09-363-573-2	Sequence 2, Appl
299	33	45.2	574	6	5223254-2	Patent No. 5223254	372	32	43.8	120	3	US-09-363-573-3	Sequence 3, Appl
300	33	45.2	594	2	US-08-659-939-2	Sequence 2, Appl	373	32	43.8	120	3	US-08-530-340-6	Sequence 6, Appl
301	33	45.2	594	2	US-08-850-041-2	Sequence 2, Appl	374	32	43.8	120	3	US-08-845-541B-1	Sequence 1, Appl
302	33	45.2	594	2	US-08-467-963C-6	Sequence 6, Appl	375	32	43.8	120	3	US-08-845-541B-3	Sequence 3, Appl
303	33	45.2	594	2	US-08-838-189D-6	Sequence 6, Appl	376	32	43.8	120	3	US-08-845-541B-4	Sequence 4, Appl
304	33	45.2	594	3	US-08-852-344D-6	Sequence 6, Appl	377	32	43.8	120	3	US-08-845-541B-5	Sequence 5, Appl
305	33	45.2	594	3	US-08-896-500-2	Sequence 2, Appl	378	32	43.8	120	3	US-08-845-541B-6	Sequence 6, Appl
306	33	45.2	594	3	US-08-476-397-2	Sequence 2, Appl	379	32	43.8	120	3	US-08-845-541B-7	Sequence 7, Appl
307	33	45.2	594	3	US-08-973-720-2	Sequence 2, Appl	380	32	43.8	120	3	US-08-845-541B-8	Sequence 8, Appl
308	33	45.2	594	3	US-08-344-639E-6	Sequence 6, Appl	381	32	43.8	120	3	US-08-845-541B-9	Sequence 9, Appl
309	33	45.2	594	3	US-09-262-927-2	Sequence 2, Appl	382	32	43.8	120	3	US-08-845-541B-10	Sequence 10, Appl
310	33	45.2	594	3	US-08-467-969A-6	Sequence 6, Appl	383	32	43.8	120	3	US-08-845-541B-11	Sequence 11, Appl
311	33	45.2	594	3	US-08-467-961A-6	Sequence 6, Appl	384	32	43.8	120	3	US-08-845-541B-12	Sequence 12, Appl
312	33	45.2	594	3	US-08-001-554A-6	Sequence 6, Appl	385	32	43.8	120	3	US-08-845-541B-13	Sequence 13, Appl
313	33	45.2	594	4	US-09-570-383-2	Sequence 2, Appl	386	32	43.8	120	3	US-08-845-541B-17	Sequence 17, Appl
314	33	45.2	594	4	US-09-395-336-2	Sequence 2, Appl	387	32	43.8	120	3	US-08-845-541B-18	Sequence 18, Appl
315	33	45.2	681	6	5194595-19	Patent No. 5194595	388	32	43.8	120	3	US-08-845-541B-19	Sequence 19, Appl
316	33	45.2	681	6	5194595-19	Patent No. 5194595	389	32	43.8	120	3	US-08-845-541B-20	Sequence 20, Appl
317	33	45.2	686	4	US-09-252-991A-18115	Sequence 18115, A	390	32	43.8	120	3	US-08-845-541B-21	Sequence 21, Appl
318	33	45.2	730	1	US-08-121-713D-58	Sequence 58, Appl	391	32	43.8	120	3	US-09-066-065A-1	Sequence 1, Appl
319	33	45.2	730	1	US-08-835-268-58	Sequence 58, Appl	392	32	43.8	120	3	US-09-066-065A-3	Sequence 3, Appl


```
; Patent No. 6583266
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; APPLICANT: MAO, JEN-I
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P. C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,731A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: GATES, EDWARD R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: C0044/7125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/720-3500
; TELEFAX: 617/720-2441
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 705 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: MYCOBACTERIUM TUBERCULOSIS
; US-08-311-731A-4

Query Match 60.3%; Score 44; DB 4; Length 705;
Best Local Similarity 70.0%; Pred. No. 27;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 YRCRGDDSKV 11
Db 383 YRCRADTEV 392

RESULT 3
US-09-360-376-54
; Sequence 54, Application US/09360376
; Patent No. 6495739
; GENERAL INFORMATION:
; APPLICANT: Lassner, Michael
; APPLICANT: Ruezinsky, Diane
; TITLE OF INVENTION: PLANT PHOSPHATIDIC ACID PHOSPHATASES
; FILE REFERENCE: 17026/01/US
; CURRENT APPLICATION NUMBER: US/09/360,376
; CURRENT FILING DATE: 1999-07-23
; PRIOR APPLICATION NUMBER: US 09/122,315
; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 54
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Rattus sp.
; US-09-360-376-54

Query Match 54.8%; Score 40; DB 4; Length 282;
Best Local Similarity 60.0%; Pred. No. 90;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

US-09-489-039A-12030
; Sequence 12030, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12030
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-12030

Query Match 54.8%; Score 40; DB 4; Length 424;
Best Local Similarity 85.7%; Pred. No. 76;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDD 8
Db 63 PRCRGDD 69

RESULT 5
US-09-270-767-41177
; Sequence 41177, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41177
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
; US-09-270-767-41177

Query Match 54.8%; Score 40; DB 4; Length 501;
Best Local Similarity 60.0%; Pred. No. 90;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 YRCRGDDSKV 11
Db 485 YRCRGDDSLI 494

RESULT 6
US-09-270-767-56393
; Sequence 56393, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
```

```

; FILE REFERENCE: File Reference: 7336-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 56393
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-56393

```

Query Match 54.8%; Score 40; DB 4; Length 501;
Best Local Similarity 60.0%; Pred. No. 90;
Matches 6; Conservative 2; Mismatches 2; Indels

```

QY      2 YRCRGDDSKV 11
      |||:| | | :
Db     485 YRCKGDSSLI 494

```

RESULT 7
US-08-992-035A-3
Sequence 3, Application US/08992035A
Patent No. 6242179
GENERAL INFORMATION:
APPLICANT: Shah, Purvi
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Lal, Preeti
TITLE OF INVENTION: HUMAN PHOSPHATASES
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

Query Match 53.4%; Score 39; DB 3; Length 283;
Best Local Similarity 46.2%; Pred. No. 74;
Matches 6; Conservative 5; Mismatches 2; Indels

```
Qy      1 NYRCRGDDSKVQE 13
        :|:|:|:|:|:|:|
Db     150 DYICOGNEEKVKE 162
```

```

RESULT 8
US-09-360-376-53
; Sequence 53, Application US/09360376
; Patent NO. 6495739
; GENERAL INFORMATION:
; APPLICANT: Lassarner, Michael
; APPLICANT: Ruzinsky, Diane
; TITLE OF INVENTION: PLANT PHOSPHATIDIC ACID PHOSPHATASES
; FILE REFERENCE: 17026/01/US
; CURRENT APPLICATION NUMBER: US/09/360,376
; CURRENT FILING DATE: 1999-07-23
; PRIOR APPLICATION NUMBER: US 09/122,315
; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 53
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-360-376-53

```

Query Match	53.4%	Score 39;	DB 4;	Length 283;
Best Local Similarity	46.2%;	Pred. No. 74;		
Matches	6;	Conservative	5;	Mismatches
				2;
				Indels
				0;
				Gaps
				0;

Qy 1 NYRCRGDDSKVQE 13
:
Dp 150 DYICOGNEEKVKE 162

```

RESULT 9
US-09-513-999C-6407
; Sequence 6407, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 6407
; LENGTH: 120
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 67
; OTHER INFORMATION: Xaa=Arg or Ser
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 70
; OTHER INFORMATION: Xaa=Lys or Asn or Thr
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 94
; OTHER INFORMATION: Xaa= * or Ala or Glu or Gly or Lys or Leu or Met
; US-09-513-999C-6407

```

Query Match 52.1%; Score 38; DB 4; Length 120;
Best Local Similarity 54.5%; Pred. No. 46;
Matches 6: Conservative 3: Mismatches 2: Indels

```
QY      2 YRCRGDDSKVQ 12
      |||:| :||
Db      101 YRCCKGNVRVQ 111

RESULT 10
US-09-690-454-131
; Sequence 131, Application US/09690454
; Patent No. 6531447
; GENERAL INFORMATION:
; APPLICANT: Steven M. Ruben, et al.
; TITLE OF INVENTION: 32 Human Secreted Proteins
; FILE REFERENCE: PZ006P1
; CURRENT APPLICATION NUMBER: US/09/690,454
; CURRENT FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: 09/189,144
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 60/044,039
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,093
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,190
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/050,935
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,101
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,356
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/056,250
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,296
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,293
; PRIOR FILING DATE: August 29, 1997
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 131
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-690-454-131

Query Match      52.1%; Score 38; DB 4; Length 136;
Best Local Similarity 54.5%; Pred. No. 52;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      2 YRCRGDDSKVQ 12
      |||:| :||
Db      61 YRCCKGNVRVQ 71

RESULT 11
US-09-015-734-12
; Sequence 12, Application US/09015734
; Patent No. 6057127
; GENERAL INFORMATION:
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL EQUINE Fc EPSILON RECEPTOR ALPHA
; TITLE OF INVENTION: CHAIN NUCLEIC ACID MOLECULES, PROTEINS AND USES THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heeka Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/515,311
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/015,734
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: DI-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; COMPUTER: IBM PC compatible

US-09-015-734-12

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/015,734
; FILING DATE: 29-JAN-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: DI-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
US-09-015-734-12

Query Match      50.7%; Score 37; DB 3; Length 201;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      2 YRCRGDDSKVQ 13
      |||:| :||
Db      91 YRCNRNDLNLSE 102

RESULT 12
US-09-515-311-12
; Sequence 12, Application US/09515311
; Patent No. 6582701
; GENERAL INFORMATION:
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL EQUINE Fc EPSILON RECEPTOR ALPHA
; TITLE OF INVENTION: CHAIN NUCLEIC ACID MOLECULES, PROTEINS AND USES THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heeka Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/515,311
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/015,734
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: DI-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
US-09-515-311-12
```

```
Query Match          50.7%; Score 37; DB 4; Length 201;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 YRCRGDDSKVQE 13
   |||| :| :|
Db 91 YRCRNDLNLSE 102

RESULT 13
US-09-015-734-7
; Sequence 7, Application US/09015734
; Patent No. 6057127
; GENERAL INFORMATION:
; APPLICANT: Weber, Eric R.
; APPLICANT: McCall, Catherine A.
; TITLE OF INVENTION: NOVEL EQUINE Fc EPSILON RECEPTOR ALPHA
; CHAIN OF INVENTION: CHAIN NUCLEIC ACID MOLECULES, PROTEINS AND USES THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/015,734
; FILING DATE: 29-JAN-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: DI-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 236 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; US-09-015-734-7

Query Match          50.7%; Score 37; DB 3; Length 236;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 YRCRGDDSKVQE 13
   |||| :| :|
Db 72 YRCRNDLNLSE 83

RESULT 14
US-09-515-311-7
; Sequence 7, Application US/09515311
; Patent No. 6582701
; GENERAL INFORMATION:
; APPLICANT: Weber, Eric R.
; APPLICANT: McCall, Catherine A.
; TITLE OF INVENTION: NOVEL EQUINE Fc EPSILON RECEPTOR ALPHA
; CHAIN OF INVENTION: CHAIN NUCLEIC ACID MOLECULES, PROTEINS AND USES THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/015,734
; FILING DATE: 29-JAN-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: DI-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 236 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; US-09-015-734-7

Query Match          50.7%; Score 37; DB 4; Length 201;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 YRCRGDDSKVQE 13
   |||| :| :|
Db 91 YRCRNDLNLSE 102

RESULT 15
US-09-015-734-2
; Sequence 2, Application US/09015734
; Patent No. 6057127
; GENERAL INFORMATION:
; APPLICANT: Weber, Eric R.
; APPLICANT: McCall, Catherine A.
; TITLE OF INVENTION: NOVEL EQUINE Fc EPSILON RECEPTOR ALPHA
; CHAIN OF INVENTION: CHAIN NUCLEIC ACID MOLECULES, PROTEINS AND USES THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/015,734
; FILING DATE: 29-JAN-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: DI-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 236 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; US-09-515-311-7

Query Match          50.7%; Score 37; DB 4; Length 236;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 YRCRGDDSKVQE 13
   |||| :| :|
Db 72 YRCRNDLNLSE 83
```

```
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 255 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: Protein
US-09-015-734-2

Query Match          50.7%; Score 37; DB 3; Length 255;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      2 YRCRGDDSKVQE 13
      |||||.:|.:|
Db      91 YRCRNDNLNLS 102

RESULT 16
US-09-515-311-2
; Sequence 2, Application US/09515311
; Patent No. 6582701
; GENERAL INFORMATION:
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL EQUINE Fc EPSILON RECEPTOR ALPHA
; TITLE OF INVENTION: CHAIN NUCLEIC ACID MOLECULES, PROTEINS AND USES THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/515,311
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/015,734
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: DI-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 255 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: Protein
US-09-515-311-2

Query Match          50.7%; Score 37; DB 4; Length 255;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      2 YRCRGDDSKVQE 13
      |||||.:|.:|
Db      91 YRCRNDNLNLS 102

RESULT 17
US-09-489-039A-13748
; Sequence 13748, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13748
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13748

Query Match          50.7%; Score 37; DB 4; Length 272;
Best Local Similarity 46.2%; Pred. No. 1.5e+02;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY      1 NYRCRGDDSKVQE 13
      ||.:|||.:|
Db      210 NVRFKGGDDAKAEK 222

RESULT 18
US-09-489-039A-11789
; Sequence 11789, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11789
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11789

Query Match          50.7%; Score 37; DB 4; Length 275;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 NYRCRGDDSD 9
      |.:|.:|.:|
Db      32 NWRCGGDSDNS 40

RESULT 19
US-09-489-039A-13164
; Sequence 13164, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13164
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
```

```
US-09-489-039A-13164
Query Match          50.7%; Score 37; DB 4; Length 488;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 RCRGDD 8
DB      102 RCRGDD 107

RESULT 20
US-09-601-478-5
; Sequence 5, Application US/09601478
; Patent No. 6403785
; GENERAL INFORMATION:
; APPLICANT: Oetuka Pharmaceutical Co., Ltd.
; TITLE OF INVENTION: Human TSC403 gene and human INGIL gene
; FILE REFERENCE: Q60193
; CURRENT APPLICATION NUMBER: US/09/601,478
; CURRENT FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: JP H10-134679
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: JP H10-73234
; PRIOR FILING DATE: 1998-03-05
; PRIOR APPLICATION NUMBER: JP H10-38133
; PRIOR FILING DATE: 1998-02-03
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human embryonic brain cDNA library
US-09-601-478-5

Query Match          49.3%; Score 36; DB 4; Length 280;
Best Local Similarity 45.5%; Pred. No. 2.3e+02;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      3 RCRGDSKVQE 13
DB      257 KCRGDNEXTMD 267

RESULT 21
US-09-601-478-8
; Sequence 8, Application US/09601478
; Patent No. 6403785
; GENERAL INFORMATION:
; APPLICANT: Oetuka Pharmaceutical Co., Ltd.
; TITLE OF INVENTION: Human TSC403 gene and human INGIL gene
; FILE REFERENCE: Q60193
; CURRENT APPLICATION NUMBER: US/09/601,478
; CURRENT FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: JP H10-134679
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: JP H10-73234
; PRIOR FILING DATE: 1998-03-05
; PRIOR APPLICATION NUMBER: JP H10-38133
; PRIOR FILING DATE: 1998-02-03
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: human embryonic brain cDNA library
US-09-601-478-8

Query Match          49.3%; Score 36; DB 4; Length 280;
Best Local Similarity 45.5%; Pred. No. 2.3e+02;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

US-09-442-013-13
; Sequence 13, Application US/09442013
; Patent No. 6737232
; GENERAL INFORMATION:
; APPLICANT: Lou, Ying
; APPLICANT: Xu, Xiang
; APPLICANT: Leo, Cindy
; APPLICANT: Huang, Betty
; APPLICANT: Shen, Mary
; TITLE OF INVENTION: NOVEL IAPS ASSOCIATED CELL CYCLE PROTEINS, COMPOSITIONS
; FILE REFERENCE: A-68289/DJB/RMS/DAV
; CURRENT APPLICATION NUMBER: US/09/442,013
; CURRENT FILING DATE: 1999-11-17
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION:
US-09-442-013-13

Query Match          49.3%; Score 36; DB 4; Length 280;
Best Local Similarity 45.5%; Pred. No. 2.3e+02;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      3 RCRGDSKVQE 13
DB      257 KCRGDNEXTMD 267

RESULT 23
US-09-513-365A-1
; Sequence 1, Application US/09513365A
; Patent No. 6790948
; GENERAL INFORMATION:
; APPLICANT: Harris, Curtis C
; APPLICANT: Nagashima, Makoto
; APPLICANT: Government of United States as represented by the Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: New Tumor Suppressor Gene P33ING2
; FILE REFERENCE: 015280-376100US
; CURRENT APPLICATION NUMBER: US/09/513,365A
; CURRENT FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: US 60/121,891
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: p33ING2
US-09-513-365A-1

Query Match          49.3%; Score 36; DB 4; Length 280;
Best Local Similarity 45.5%; Pred. No. 2.3e+02;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      3 RCRGDSKVQE 13
DB      257 KCRGDNEXTMD 267
```



```
RESULT 24
US-09-513-365A-6
; Sequence 6, Application US/09513365A
; Patent No. 6790948
; GENERAL INFORMATION:
; APPLICANT: Harris, Curtis C
; APPLICANT: Nagashima, Makoto
; APPLICANT: Government of United States as represented by the Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: New Tumor Suppressor Gene P33ING2
; FILE REFERENCE: 015280-376100US
; CURRENT APPLICATION NUMBER: US/09/513,365A
; CURRENT FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: US 60/121,891
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: missense
; OTHER INFORMATION: P33ING2 sequence - Arg 153 to Ser
US-09-513-365A-6

Query Match          49.3%; Score 36; DB 4; Length 280;
Best Local Similarity 45.5%; Pred. No. 2.3e+02;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      3 RCGRGDDSKVQ 13
Db      257 KCRGDNKTKMD 267

RESULT 25
US-09-328-352-8163
; Sequence 8163, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 8163
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-8163

Query Match          49.3%; Score 36; DB 4; Length 316;
Best Local Similarity 63.6%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      2 YRCRGDDSKVQ 12
Db      175 YRCRGTYYKAQ 185

RESULT 26
US-09-949-016-8237
; Sequence 8237, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
```

```
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8237
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8237

Query Match          49.3%; Score 36; DB 4; Length 416;
Best Local Similarity 60.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 NYRCRGDDSK 10
Db      400 NWRCPGTDSE 409

RESULT 27
US-09-786-240-1
; Sequence 1, Application US/09786240
; Patent No. 6558935
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: TANG, Y. Tom
; APPLICANT: CORLEY, Neil C.
; APPLICANT: GUEGLER, Karl J.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LAL, Preeti
; APPLICANT: YUE, Henry
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: AZIMZAI, Yalda
; TITLE OF INVENTION: HUMAN TRANSFERASE PROTEINS
; FILE REFERENCE: PF-0592 PCT
; CURRENT APPLICATION NUMBER: US/09/786,240
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/150,657; unassigned; 09/186,779; unassigned; 60/133,642
; PRIOR FILING DATE: 1998-09-10; 1998-09-10; 1998-11-04; 1998-11-04; 1999-05-11
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6558935 1632930CD1
US-09-786-240-1

Query Match          49.3%; Score 36; DB 4; Length 498;
Best Local Similarity 60.0%; Pred. No. 4.1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 NYRCRGDDSK 10
Db      482 NWRCPGTDSE 491

RESULT 28
US-09-198-452A-1290
; Sequence 1290, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffois, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevent
; TITLE OF INVENTION: and treatment of infection
```

FILE REFERENCE: 9710-003-999
CURRENT APPLICATION NUMBER: US/09/198,452A
CURRENT FILING DATE: 1998-11-24
NUMBER OF SEQ ID NOS: 6849
SEQ ID NO 1290
LENGTH: 94
TYPE: PRT
ORGANISM: Chlamydia pneumoniae
US-09-198-452A-1290

Query Match 47.9%; Score 35; DB 4; Length 94;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVRCRGD 7
| | | | |
DB 13 NRRCRGD 19

RESULT 29
US-08-473-981A-12
Sequence 12, Application US/08473981A
Patent No. 5629162
GENERAL INFORMATION:
APPLICANT: deFougerolles, Antonin R
APPLICANT: Springer, Timothy A
TITLE OF INVENTION: METHODS OF IDENTIFYING AGENTS WHICH MODULATE
TITLE OF INVENTION: ICAM-3 BINDING TO LFA-1 (AS AMENDED)
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
STREET: 1100 NEW YORK AVENUE, N.W. SUITE 600
CITY: WASHINGTON
STATE: D. C.
COUNTRY: USA
ZIP: 20005

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/473,981A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
NAME: MILLONIG, ROBERT C
REGISTRATION NUMBER: 34,395
REFERENCE/DOCKET NUMBER: 1011.0560004
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 99 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-473-981A-12

Query Match 47.9%; Score 35; DB 1; Length 99;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 CRGDDSKVOE 13
| | | | |
DB 19 CPGSDEKVEE 28

RESULT 30
US-08-474-087-12
Sequence 12, Application US/08474087

Patent No. 5891841
GENERAL INFORMATION:
APPLICANT: de Fougerolles, Antonin R
APPLICANT: Springer, Timothy A
TITLE OF INVENTION: METHODS OF USING INTERCELLULAR ADHESION MOLECULE-
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
STREET: 1100 NEW YORK AVENUE, N.W. SUITE 600
CITY: WASHINGTON
STATE: D. C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,087
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/038,990
FILING DATE: 23-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/712,879
FILING DATE: 11-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: MILLONIG, ROBERT C
REGISTRATION NUMBER: 34,395
REFERENCE/DOCKET NUMBER: 1011.0560003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 99 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-474-087-12

Query Match 47.9%; Score 35; DB 2; Length 99;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 CRGDDSKVOE 13
| | | | |
DB 19 CPGSDEKVEE 28

RESULT 31
US-09-252-991A-22600
Sequence 22600, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 22600
LENGTH: 213
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-22600

Query Match 47.9%; Score 35; DB 4; Length 213;
Best Local Similarity 46.2%; Pred. No. 2.6e+02;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 NYRCRGDDSKVQE 13
Db 84 DYTQRGDERKVED 96

RESULT 32

US-10-000-489-78

; Sequence 78, Application US/10000489
; Patent No. 6794363
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US6.DIV
; CURRENT APPLICATION NUMBER: US/10/000,489
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: Jpatent
; SEQ ID NO 78
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: 1..21
US-10-000-489-78

Query Match 47.9%; Score 35; DB 4; Length 269;
Best Local Similarity 66.7%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 YRCRGDDSK 10
Db 183 YRCNGTSSK 191

RESULT 33

US-09-949-016-6121
; Sequence 6121, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6121
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6121

Query Match 47.9%; Score 35; DB 4; Length 269;
Best Local Similarity 66.7%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 YRCRGDDSK 10
Db 183 YRCNGTSSK 191

RESULT 34

US-08-282-951-6
; Sequence 6, Application US/08282951
; Patent No. 5685590
; GENERAL INFORMATION:
; APPLICANT: YANG, ZHI
; TITLE OF INVENTION: METHOD FOR ISOLATING AND DIRECTLY
; TITLE OF INVENTION: CLONING GENES WHICH ENCODE CELL-SURFACE AND SECRETED
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/282,951
; FILING DATE: 29-JUL-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: LEHNHARDT, SUSAN K.
; REGISTRATION NUMBER: 33,943
; REFERENCE/DOCKET NUMBER: 20296-20012.00
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 272 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-282-951-6

Query Match 47.9%; Score 35; DB 1; Length 272;
Best Local Similarity 54.5%; Pred. No. 3.3e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 YRCRGDDSKVQ 12
Db 91 YRCQATDAKGQ 101

RESULT 35

US-09-949-016-7261
; Sequence 7261, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

```
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 7261
/ LENGTH: 276
/ TYPE: PRT
/ ORGANISM: Human
/ US-09-949-016-7261

Query Match 47.9%; Score 35; DB 4; Length 276;
Best Local Similarity 66.7%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 YRCRGDSSK 10
Db 190 YRCNGTSSK 198

RESULT 36
US-08-992-035A-1
/ Sequence 1, Application US/08992035A
/ Patent No. 6242179
/ GENERAL INFORMATION:
/ APPLICANT: Shah, Purvi
/ APPLICANT: Hillman, Jennifer L.
/ APPLICANT: Corley, Neil C.
/ APPLICANT: Lal, Preeti
/ TITLE OF INVENTION: HUMAN PHOSPHATASES
/ NUMBER OF SEQUENCES: 6
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Incyte Pharmaceuticals, Inc.
/ STREET: 3174 Porter Dr.
/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94304
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/992,035A
/ FILING DATE: December 17, 1997
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Billings, Lucy J.
/ REGISTRATION NUMBER: 36,749
/ REFERENCE/DOCKET NUMBER: PF-0433 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 650-855-0555
/ TELEFAX: 650-845-4166
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 285 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: BLADNOT06
/ CLONE: 1719418
/ US-08-992-035A-1
```

```
Query Match 47.9%; Score 35; DB 3; Length 285;
Best Local Similarity 50.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 YRCRGDDSKVQE 13
Db 152 YICRGNAERVKE 163

RESULT 37
US-09-506-286B-92
/ Sequence 92, Application US/09506286B
/ Patent No. 6482414
/ GENERAL INFORMATION:
/ APPLICANT: Dowling, Patricia W.
/ APPLICANT: Youngner, Julius S.
/ APPLICANT: The University of Pittsburgh, of the Commonwealth
/ TITLE OF INVENTION: COLD-ADAPTED EQUINE INFLUENZA VIRUSES
/ FILE REFERENCE: EQ-1-C2
/ CURRENT APPLICATION NUMBER: US/09/506,286B
/ CURRENT FILING DATE: 2000-02-16
/ PRIOR APPLICATION NUMBER: 09/133,921
/ PRIOR FILING DATE: 1998-08-13
/ PRIOR APPLICATION NUMBER: PCT/US99/18583
/ PRIOR FILING DATE: 1999-08-12
/ NUMBER OF SEQ ID NOS: 108
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 92
/ LENGTH: 396
/ TYPE: PRT
/ ORGANISM: Equine influenza virus H3N8
/ US-09-506-286B-92

Query Match 47.9%; Score 35; DB 4; Length 396;
Best Local Similarity 45.5%; Pred. No. 4.8e+02;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 YRCRGDDSKVQ 12
Db 198 YRCQRGDTQIQ 208

RESULT 38
US-09-506-286B-95
/ Sequence 95, Application US/09506286B
/ Patent No. 6482414
/ GENERAL INFORMATION:
/ APPLICANT: Dowling, Patricia W.
/ APPLICANT: Youngner, Julius S.
/ APPLICANT: The University of Pittsburgh, of the Commonwealth
/ TITLE OF INVENTION: COLD-ADAPTED EQUINE INFLUENZA VIRUSES
/ FILE REFERENCE: EQ-1-C2
/ CURRENT APPLICATION NUMBER: US/09/506,286B
/ CURRENT FILING DATE: 2000-02-16
/ PRIOR APPLICATION NUMBER: 09/133,921
/ PRIOR FILING DATE: 1998-08-13
/ PRIOR APPLICATION NUMBER: PCT/US99/18583
/ PRIOR FILING DATE: 1999-08-12
/ NUMBER OF SEQ ID NOS: 108
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 95
/ LENGTH: 396
/ TYPE: PRT
/ ORGANISM: Equine influenza virus H3N8
/ US-09-506-286B-95

Query Match 47.9%; Score 35; DB 4; Length 396;
Best Local Similarity 45.5%; Pred. No. 4.8e+02;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 YRCRGDDSKVQ 12
Db 198 YRCQRGDTQIQ 208
```

```
Db      198 YRCRGDTQIQ 208

RESULT 39
US-10-065-133A-92
; Sequence 92, Application US/10065133A
; Patent No. 6685946
; GENERAL INFORMATION:
; APPLICANT: Dowling, Patricia W.
; APPLICANT: Youngner, Julius S.
; TITLE OF INVENTION: COLD-ADAPTED EQUINE INFLUENZA VIRUSES
; FILE REFERENCE: EQ-1-C2-1
; CURRENT APPLICATION NUMBER: US/10/065,133A
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: PCT/US99/18583
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 09/133,921
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 92
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Equine influenza virus H3N8
US-10-065-133A-92

Query Match      47.9%; Score 35; DB 4; Length 396;
Best Local Similarity 45.5%; Pred. No. 4.8e+02;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY      2 YRCRGDSKVQ 12
      |||: |||:
Db      198 YRCRGDTQIQ 208

RESULT 40
US-10-065-133A-95
; Sequence 95, Application US/10065133A
; Patent No. 6685946
; GENERAL INFORMATION:
; APPLICANT: Dowling, Patricia W.
; APPLICANT: Youngner, Julius S.
; TITLE OF INVENTION: COLD-ADAPTED EQUINE INFLUENZA VIRUSES
; FILE REFERENCE: EQ-1-C2-1
; CURRENT APPLICATION NUMBER: US/10/065,133A
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: PCT/US99/18583
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 09/133,921
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 95
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Equine influenza virus H3N8
US-10-065-133A-95

Query Match      47.9%; Score 35; DB 4; Length 396;
Best Local Similarity 45.5%; Pred. No. 4.8e+02;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY      2 YRCRGDSKVQ 12
      |||: |||:
Db      198 YRCRGDTQIQ 208

TITLE OF INVENTION: Isolated Mammalian Monocyte Cell Genes;
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESS: DNAX Research Institute
STREET: 901 California Avenue
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/985,950
FILING DATE: 05-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/041,279
FILING DATE: 21-MARCH-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/033,181
FILING DATE: 16-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/032,252
FILING DATE: 06-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: DX0670K
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650)852-9196
TELEFAX: (650)496-1204
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 439 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-985-950-12

Query Match      47.9%; Score 35; DB 3; Length 439;
Best Local Similarity 60.0%; Pred. No. 5.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 NYRCRGDSK 10
      :|||: |||:
Db      194 SYRCGYDSR 203

RESULT 42
US-09-310-463-8
; Sequence 8, Application US/09310463A
; Patent No. 6384203
; GENERAL INFORMATION:
; APPLICANT: Cosman, David J.
; APPLICANT: Anderson, Dirk M.
; APPLICANT: Borges, Luis
; TITLE OF INVENTION: Family of Immunoregulators Designated Leukocyte Immunoglobulin-
; TITLE OF INVENTION: Like Receptors (LIR)
; FILE REFERENCE: 2624-A
; CURRENT APPLICATION NUMBER: US/09/310,463A
; CURRENT FILING DATE: 1999-05-12
; EARLIER APPLICATION NUMBER: 08/842,248
; EARLIER FILING DATE: 1997-04-24
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 8
; LENGTH: 439
; TYPE: PRT
; ORGANISM: human
US-09-310-463-8
```

```

Query Match          47.9%; Score 35; DB 3; Length 439;
Best Local Similarity 60.0%; Pred. No. 5.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
DB 194 SYRCGYDSR 203

RESULT 43
US-08-842-248A-8
; Sequence 8, Application US/08842248A
; Patent No. 6448035
; GENERAL INFORMATION:
; APPLICANT: Cosman, David J.
; TITLE OF INVENTION: Family of Immunoregulators Designated
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Janis C. Henry, Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: US
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM/PC Compatible
; OPERATING SYSTEM: Microsoft Word 7.0
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/842,248A
; FILING DATE: April 24, 1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Henry, Janis C.
; REGISTRATION NUMBER: 34,347
; REFERENCE/DOCKET NUMBER: 2624
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; TELEX: 756822
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 439 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-842-248A-8

Query Match          47.9%; Score 35; DB 4; Length 439;
Best Local Similarity 60.0%; Pred. No. 5.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
DB 194 SYRCGYDSR 203

RESULT 44
US-09-546-049-12
; Sequence 12, Application US/09546049
; Patent No. 6479638
; GENERAL INFORMATION:
; APPLICANT: Adema, Gosse Jan
; Meyaard, Linde
; Gorman, Daniel M.
; McClanahan, Terrill K.
; Zurawski, Sandra M.
; Zurawski, Gerard
; Lanier, Lewis L.
; Phillips Jr., Joseph H.

```

```

; TITLE OF INVENTION: Isolated Mammalian Monocyte Cell Genes;
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/546,049
; FILING DATE: 10-Apr-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/985,950
; FILING DATE: 05-DEC-1997
; APPLICATION NUMBER: US 60/041,279
; FILING DATE: 21-MARCH-1997
; APPLICATION NUMBER: US 60/033,181
; FILING DATE: 16-DEC-1996
; APPLICATION NUMBER: US 60/032,252
; FILING DATE: 06-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: DX0670K
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)852-9196
; TELEFAX: (650)496-1204
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 439 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-546-049-12

Query Match          47.9%; Score 35; DB 4; Length 439;
Best Local Similarity 60.0%; Pred. No. 5.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
DB 194 SYRCGYDSR 203

RESULT 45
US-09-627-376-10
; Sequence 10, Application US/09627376
; Patent No. 6342385
; GENERAL INFORMATION:
; APPLICANT: Qi, Fengxia
; TITLE OF INVENTION: MUTACIN I BIOSYNTHESIS GENES AND PROTEINS
; FILE REFERENCE: UAB-17402/22
; CURRENT APPLICATION NUMBER: US/09/627,376
; CURRENT FILING DATE: 2001-05-30
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-627-376-10

Query Match          47.9%; Score 35; DB 3; Length 447;
Best Local Similarity 50.0%; Pred. No. 5.4e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```

QY 1 NYRCRGDSSKQV 12
|| :||| :
Db 410 NYHLOGDSSKAR 421

RESULT 46
US-10-047-676B-10
; Sequence 10, Application US/10047676B
; Patent No. 6699970
; GENERAL INFORMATION:
; APPLICANT: Qi, Fengxia
; APPLICANT: Caulfield, Page W.
; APPLICANT: Chen, Ping
; TITLE OF INVENTION: MUTACIN I BIOSYNTHESIS GENES AND PROTEINS
; FILE REFERENCE: UAB-17403/22
; CURRENT APPLICATION NUMBER: US/10/047,676B
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 09/627,376
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 10
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-047-676B-10

Query Match 47.9%; Score 35; DB 4; Length 447;
Best Local Similarity 50.0%; Pred. No. 5.4e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 NYRCRGDSSKQV 12
|| :||| :
Db 410 NYHLOGDSSKAR 421

RESULT 47
US-09-248-796A-17110
; Sequence 17110, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17110
; LENGTH: 503
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-17110

Query Match 47.9%; Score 35; DB 4; Length 503;
Best Local Similarity 66.7%; Pred. No. 6.1e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 YRCRGDSSK 10
||| :||| :
Db 288 YRCFSDDDK 296

RESULT 48
US-08-419-652-4
; Sequence 4, Application US/08419652
; Patent No. 5831007
; GENERAL INFORMATION:

; APPLICANT: Chua, Anne O
; APPLICANT: Gubler, Ulrich A
; TITLE OF INVENTION: INTERLEUKIN-12 RECEPTOR
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingeland Street
; CITY: Nutley
; STATE: New Jersey
; COUNTRY: United States of America
; ZIP: 07110-1199
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/419,652
; FILING DATE: 11-APR-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/248,532
; FILING DATE: 31-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/094,713
; FILING DATE: 19-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Kass, Alan P
; REGISTRATION NUMBER: 32142
; REFERENCE/DOCKET NUMBER: CD 9174
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 235-4205
; TELEFAX: (201) 235-3500
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 621 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Region
; LOCATION: 1..621
; OTHER INFORMATION: /note= "Represents residues 42 to
; OTHER INFORMATION: 662 of SEQ ID NO:2."
US-08-419-652-4

Query Match 47.9%; Score 35; DB 2; Length 621;
Best Local Similarity 75.0%; Pred. No. 7.6e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 RCRGDDSK 10
||| :||| :
Db 443 RCRDEDSK 450

RESULT 49
US-08-248-532-3
; Sequence 3, Application US/08248532
; Patent No. 5536657
; GENERAL INFORMATION:
; APPLICANT: Chua, Anne O
; APPLICANT: Gubler, Ulrich A
; TITLE OF INVENTION: INTERLEUKIN-12 RECEPTOR
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingeland Street
; CITY: Nutley
; STATE: New Jersey
; COUNTRY: United States of America
; ZIP: 07110-1199
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/248,532
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/094,713
FILING DATE: 19-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Kass, Alan P
REGISTRATION NUMBER: 32142
REFERENCE/DOCKET NUMBER: CD 8973
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201) 235-4205
TELEFAX: (201) 235-3500
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 660 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-248-532-3

Query Match 47.9%; Score 35; DB 1; Length 660;
Best Local Similarity 75.0%; Pred. No. 8.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 RCRGDDSK 10
||| :|||
DB 484 RCRDEDSK 491

RESULT 50
US-08-419-652-3
Sequence 3, Application US/08419652
Patent No. 5831007
GENERAL INFORMATION:
APPLICANT: Chuva, Anne O
APPLICANT: Gubler, Ulrich A
TITLE OF INVENTION: INTERLEUKIN-12 RECEPTOR
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: New Jersey
COUNTRY: United States of America
ZIP: 07110-1199
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/419,652
FILING DATE: 11-APR-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/248,532
FILING DATE: 31-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/094,713
FILING DATE: 19-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Kass, Alan P
REGISTRATION NUMBER: 32142
REFERENCE/DOCKET NUMBER: CD 9174
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201) 235-4205
TELEFAX: (201) 235-3500
INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 660 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-419-652-3
Query Match 47.9%; Score 35; DB 2; Length 660;
Best Local Similarity 75.0%; Pred. No. 8.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 3 RCRGDDSK 10
||| :|||
DB 484 RCRDEDSK 491
Search completed: September 7, 2005, 19:58:09
Job time : 17.3256 secs

1	73	100.0	13	17	US-10-812-238A-23	Sequence 23, Appl
2	73	100.0	20	17	US-10-812-238A-2	Sequence 2, Appl
3	73	100.0	311	16	US-10-655-601-2	Sequence 2, Appl
4	73	100.0	311	17	US-10-812-238A-13	Sequence 13, Appl
5	59	80.8	10	17	US-10-812-238A-10	Sequence 20, Appl
6	55	75.3	10	17	US-10-812-238A-22	Sequence 22, Appl
7	53	72.6	10	17	US-10-812-238A-21	Sequence 21, Appl
8	46	63.0	74	16	US-10-425-115-262985	Sequence 262985,
9	41	56.2	208	16	US-10-437-963-135805	Sequence 135805,
10	41	56.2	237	15	US-10-291-172-610	Sequence 610, App
11	41	56.2	237	15	US-10-221-278-610	Sequence 610, App

85	36	49.3	215	10	US-09-759-130B-123	Sequence 123, App	158	36	49.3	225	14	US-10-180-544-358
86	36	49.3	215	16	US-10-741-790-123	Sequence 123, App	159	36	49.3	225	14	US-10-180-546-358
87	36	49.3	223	16	US-10-425-115-294642	Sequence 294642,	160	36	49.3	225	14	US-10-180-547-358
88	36	49.3	225	9	US-09-731-872-393	Sequence 393, App	161	36	49.3	225	14	US-10-180-549-358
89	36	49.3	225	9	US-09-731-872-405	Sequence 405, App	162	36	49.3	225	14	US-10-180-555-358
90	36	49.3	225	9	US-09-949-192-37	Sequence 37, App	163	36	49.3	225	14	US-10-180-559-358
91	36	49.3	225	10	US-09-759-130B-132	Sequence 132, App	164	36	49.3	225	14	US-10-181-000-358
92	36	49.3	225	10	US-09-946-374-328	Sequence 328, App	165	36	49.3	225	14	US-10-183-010-358
93	36	49.3	225	10	US-09-876-997-393	Sequence 393, App	166	36	49.3	225	14	US-10-183-012-358
94	36	49.3	225	10	US-09-876-997-405	Sequence 405, App	167	36	49.3	225	14	US-10-184-614-358
95	36	49.3	225	13	US-10-006-867-120	Sequence 120, App	168	36	49.3	225	14	US-10-184-623-358
96	36	49.3	225	13	US-10-052-586-358	Sequence 358, App	169	36	49.3	225	14	US-10-184-635-358
97	36	49.3	225	13	US-10-063-547-120	Sequence 120, App	170	36	49.3	225	14	US-10-184-637-358
98	36	49.3	225	13	US-10-063-551-120	Sequence 120, App	171	36	49.3	225	14	US-10-184-646-358
99	36	49.3	225	14	US-10-174-590-358	Sequence 358, App	172	36	49.3	225	14	US-10-184-647-358
100	36	49.3	225	14	US-10-176-758-358	Sequence 358, App	173	36	49.3	225	14	US-10-184-652-358
101	36	49.3	225	14	US-10-175-737-358	Sequence 358, App	174	36	49.3	225	14	US-10-187-594-358
102	36	49.3	225	14	US-10-063-569-120	Sequence 120, App	175	36	49.3	225	14	US-10-187-596-358
103	36	49.3	225	14	US-10-174-581-358	Sequence 358, App	176	36	49.3	225	14	US-10-187-745-358
104	36	49.3	225	14	US-10-176-483-358	Sequence 358, App	177	36	49.3	225	14	US-10-187-885-358
105	36	49.3	225	14	US-10-176-749-358	Sequence 358, App	178	36	49.3	225	14	US-10-187-886-358
106	36	49.3	225	14	US-10-176-914-358	Sequence 358, App	179	36	49.3	225	14	US-10-199-464-358
107	36	49.3	225	14	US-10-176-915-358	Sequence 358, App	180	36	49.3	225	14	US-10-196-756-358
108	36	49.3	225	14	US-10-063-569-120	Sequence 120, App	181	36	49.3	225	14	US-10-176-751-358
109	36	49.3	225	14	US-10-063-513-120	Sequence 120, App	182	36	49.3	225	14	US-10-176-760-358
110	36	49.3	225	14	US-10-063-515-120	Sequence 120, App	183	36	49.3	225	14	US-10-176-990-358
111	36	49.3	225	14	US-10-063-512-120	Sequence 120, App	184	36	49.3	225	14	US-10-180-541-358
112	36	49.3	225	14	US-10-173-706-358	Sequence 358, App	185	36	49.3	225	14	US-10-180-542-358
113	36	49.3	225	14	US-10-175-728-358	Sequence 358, App	186	36	49.3	225	14	US-10-180-548-358
114	36	49.3	225	14	US-10-175-737-358	Sequence 358, App	187	36	49.3	225	14	US-10-180-551-358
115	36	49.3	225	14	US-10-175-752-358	Sequence 358, App	188	36	49.3	225	14	US-10-180-998-358
116	36	49.3	225	14	US-10-176-482-358	Sequence 358, App	189	36	49.3	225	14	US-10-180-999-358
117	36	49.3	225	14	US-10-176-757-358	Sequence 358, App	190	36	49.3	225	14	US-10-183-013-358
118	36	49.3	225	14	US-10-176-913-358	Sequence 358, App	191	36	49.3	225	14	US-10-184-612-358
119	36	49.3	225	14	US-10-180-552-358	Sequence 358, App	192	36	49.3	225	14	US-10-184-616-358
120	36	49.3	225	14	US-10-180-557-358	Sequence 358, App	193	36	49.3	225	14	US-10-184-617-358
121	36	49.3	225	14	US-10-063-502-120	Sequence 120, App	194	36	49.3	225	14	US-10-184-622-358
122	36	49.3	225	14	US-10-173-700-358	Sequence 358, App	195	36	49.3	225	14	US-10-184-628-358
123	36	49.3	225	14	US-10-174-572-358	Sequence 358, App	196	36	49.3	225	14	US-10-184-629-358
124	36	49.3	225	14	US-10-174-579-358	Sequence 358, App	197	36	49.3	225	14	US-10-184-630-358
125	36	49.3	225	14	US-10-174-582-358	Sequence 358, App	198	36	49.3	225	14	US-10-184-631-358
126	36	49.3	225	14	US-10-174-588-358	Sequence 358, App	199	36	49.3	225	14	US-10-184-632-358
127	36	49.3	225	14	US-10-175-739-358	Sequence 358, App	200	36	49.3	225	14	US-10-184-636-358
128	36	49.3	225	14	US-10-175-740-358	Sequence 358, App	201	36	49.3	225	14	US-10-184-640-358
129	36	49.3	225	14	US-10-175-743-358	Sequence 358, App	202	36	49.3	225	14	US-10-184-650-358
130	36	49.3	225	14	US-10-176-488-358	Sequence 358, App	203	36	49.3	225	14	US-10-184-651-358
131	36	49.3	225	14	US-10-176-492-358	Sequence 358, App	204	36	49.3	225	14	US-10-187-588-358
132	36	49.3	225	14	US-10-176-747-358	Sequence 358, App	205	36	49.3	225	14	US-10-187-597-358
133	36	49.3	225	14	US-10-176-750-358	Sequence 358, App	206	36	49.3	225	14	US-10-187-598-358
134	36	49.3	225	14	US-10-176-985-358	Sequence 358, App	207	36	49.3	225	14	US-10-187-600-358
135	36	49.3	225	14	US-10-176-987-358	Sequence 358, App	208	36	49.3	225	14	US-10-187-601-358
136	36	49.3	225	14	US-10-176-992-358	Sequence 358, App	209	36	49.3	225	14	US-10-187-602-358
137	36	49.3	225	14	US-10-176-993-358	Sequence 358, App	210	36	49.3	225	14	US-10-187-603-358
138	36	49.3	225	14	US-10-184-658-358	Sequence 358, App	211	36	49.3	225	14	US-10-187-741-358
139	36	49.3	225	14	US-10-176-991-358	Sequence 358, App	212	36	49.3	225	14	US-10-187-743-358
140	36	49.3	225	14	US-10-063-549-120	Sequence 120, App	213	36	49.3	225	14	US-10-187-746-358
141	36	49.3	225	14	US-10-173-695-358	Sequence 358, App	214	36	49.3	225	14	US-10-187-747-358
142	36	49.3	225	14	US-10-173-697-358	Sequence 358, App	215	36	49.3	225	14	US-10-187-751-358
143	36	49.3	225	14	US-10-173-705-358	Sequence 358, App	216	36	49.3	225	14	US-10-187-753-358
144	36	49.3	225	14	US-10-174-576-358	Sequence 358, App	217	36	49.3	225	14	US-10-187-754-358
145	36	49.3	225	14	US-10-174-585-358	Sequence 358, App	218	36	49.3	225	14	US-10-187-757-358
146	36	49.3	225	14	US-10-174-586-358	Sequence 358, App	219	36	49.3	225	14	US-10-187-884-358
147	36	49.3	225	14	US-10-175-477-358	Sequence 358, App	220	36	49.3	225	14	US-10-188-767-358
148	36	49.3	225	14	US-10-176-481-358	Sequence 358, App	221	36	49.3	225	14	US-10-188-769-358
149	36	49.3	225	14	US-10-176-485-358	Sequence 358, App	222	36	49.3	225	14	US-10-188-770-358
150	36	49.3	225	14	US-10-176-487-358	Sequence 358, App	223	36	49.3	225	14	US-10-188-773-358
151	36	49.3	225	14	US-10-176-493-358	Sequence 358, App	224	36	49.3	225	14	US-10-188-781-358
152	36	49.3	225	14	US-10-176-511-358	Sequence 358, App	225	36	49.3	225	14	US-10-194-361-358
153	36	49.3	225	14	US-10-176-911-358	Sequence 358, App	226	36	49.3	225	14	US-10-194-423-358
154	36	49.3	225	14	US-10-176-919-358	Sequence 358, App	227	36	49.3	225	14	US-10-195-897-358
155	36	49.3	225	14	US-10-176-925-358	Sequence 358, App	228	36	49.3	225	14	US-10-195-901-358
156	36	49.3	225	14	US-10-176-978-358	Sequence 358, App	229	36	49.3	225	14	US-10-195-902-358
157	36	49.3	225	14	US-10-179-510-358	Sequence 358, App	230	36	49.3	225	14	US-10-196-743-358
158	36	49.3	225	14	US-10-180-543-358	Sequence 358, App	231	36	49.3	225	14	US-10-196-743-358

377	36	49.3	225	14	US-10-194-363-358	Sequence 358, App	450	36	49.3	225	14	US-10-199-312-358	Sequence 358, App
378	36	49.3	225	14	US-10-194-460-358	Sequence 358, App	451	36	49.3	225	14	US-10-199-315-358	Sequence 358, App
379	36	49.3	225	14	US-10-194-463-358	Sequence 358, App	452	36	49.3	225	14	US-10-199-316-358	Sequence 358, App
380	36	49.3	225	14	US-10-194-484-358	Sequence 358, App	453	36	49.3	225	14	US-10-199-457-358	Sequence 358, App
381	36	49.3	225	14	US-10-198-884-358	Sequence 358, App	454	36	49.3	225	14	US-10-199-459-358	Sequence 358, App
382	36	49.3	225	14	US-10-196-744-358	Sequence 358, App	455	36	49.3	225	14	US-10-199-460-358	Sequence 358, App
383	36	49.3	225	14	US-10-196-755-358	Sequence 358, App	456	36	49.3	225	14	US-10-199-461-358	Sequence 358, App
384	36	49.3	225	14	US-10-197-704-358	Sequence 358, App	457	36	49.3	225	14	US-10-199-667-358	Sequence 358, App
385	36	49.3	225	14	US-10-197-710-358	Sequence 358, App	458	36	49.3	225	14	US-10-199-673-358	Sequence 358, App
386	36	49.3	225	14	US-10-198-758-358	Sequence 358, App	459	36	49.3	225	14	US-10-201-321-358	Sequence 358, App
387	36	49.3	225	14	US-10-198-766-358	Sequence 358, App	460	36	49.3	225	14	US-10-201-322-358	Sequence 358, App
388	36	49.3	225	14	US-10-199-304-358	Sequence 358, App	461	36	49.3	225	14	US-10-201-326-358	Sequence 358, App
389	36	49.3	225	14	US-10-199-309-358	Sequence 358, App	462	36	49.3	225	14	US-10-201-532-358	Sequence 358, App
390	36	49.3	225	14	US-10-199-313-358	Sequence 358, App	463	36	49.3	225	14	US-10-201-533-358	Sequence 358, App
391	36	49.3	225	14	US-10-199-456-358	Sequence 358, App	464	36	49.3	225	14	US-10-201-535-358	Sequence 358, App
392	36	49.3	225	14	US-10-201-329-358	Sequence 358, App	465	36	49.3	225	14	US-10-201-769-358	Sequence 358, App
393	36	49.3	225	14	US-10-202-412-358	Sequence 358, App	466	36	49.3	225	14	US-10-201-771-358	Sequence 358, App
394	36	49.3	225	14	US-10-206-919-358	Sequence 358, App	467	36	49.3	225	14	US-10-201-854-358	Sequence 358, App
395	36	49.3	225	14	US-10-206-922-358	Sequence 358, App	468	36	49.3	225	14	US-10-202-410-358	Sequence 358, App
396	36	49.3	225	14	US-10-206-924-358	Sequence 358, App	469	36	49.3	225	14	US-10-202-473-358	Sequence 358, App
397	36	49.3	225	14	US-10-206-928-358	Sequence 358, App	470	36	49.3	225	14	US-10-202-474-358	Sequence 358, App
398	36	49.3	225	14	US-10-207-914-358	Sequence 358, App	471	36	49.3	225	14	US-10-205-503-358	Sequence 358, App
399	36	49.3	225	14	US-10-207-921-358	Sequence 358, App	472	36	49.3	225	14	US-10-205-512-358	Sequence 358, App
400	36	49.3	225	14	US-10-207-922-358	Sequence 358, App	473	36	49.3	225	14	US-10-205-892-358	Sequence 358, App
401	36	49.3	225	14	US-10-208-027-358	Sequence 358, App	474	36	49.3	225	14	US-10-205-894-358	Sequence 358, App
402	36	49.3	225	14	US-10-013-907A-328	Sequence 328, App	475	36	49.3	225	14	US-10-205-896-358	Sequence 358, App
403	36	49.3	225	14	US-10-015-499A-328	Sequence 328, App	476	36	49.3	225	14	US-10-205-898-358	Sequence 358, App
404	36	49.3	225	14	US-10-063-555-120	Sequence 120, App	477	36	49.3	225	14	US-10-205-901-358	Sequence 358, App
405	36	49.3	225	14	US-10-196-757-358	Sequence 358, App	478	36	49.3	225	14	US-10-205-903-358	Sequence 358, App
406	36	49.3	225	14	US-10-196-594-120	Sequence 120, App	479	36	49.3	225	14	US-10-206-909-358	Sequence 358, App
407	36	49.3	225	14	US-10-196-754-358	Sequence 358, App	480	36	49.3	225	14	US-10-206-910-358	Sequence 358, App
408	36	49.3	225	14	US-10-174-571-358	Sequence 358, App	481	36	49.3	225	14	US-10-206-911-358	Sequence 358, App
409	36	49.3	225	14	US-10-176-746-358	Sequence 358, App	482	36	49.3	225	14	US-10-206-912-358	Sequence 358, App
410	36	49.3	225	14	US-10-176-923-358	Sequence 358, App	483	36	49.3	225	14	US-10-206-913-358	Sequence 358, App
411	36	49.3	225	14	US-10-183-011-358	Sequence 358, App	484	36	49.3	225	14	US-10-206-914-358	Sequence 358, App
412	36	49.3	225	14	US-10-184-633-358	Sequence 358, App	485	36	49.3	225	14	US-10-206-920-358	Sequence 358, App
413	36	49.3	225	14	US-10-184-639-358	Sequence 358, App	486	36	49.3	225	14	US-10-206-921-358	Sequence 358, App
414	36	49.3	225	14	US-10-187-742-358	Sequence 358, App	487	36	49.3	225	14	US-10-206-923-358	Sequence 358, App
415	36	49.3	225	14	US-10-187-748-358	Sequence 358, App	488	36	49.3	225	14	US-10-206-925-358	Sequence 358, App
416	36	49.3	225	14	US-10-188-766-358	Sequence 358, App	489	36	49.3	225	14	US-10-206-926-358	Sequence 358, App
417	36	49.3	225	14	US-10-188-771-358	Sequence 358, App	490	36	49.3	225	14	US-10-206-927-358	Sequence 358, App
418	36	49.3	225	14	US-10-192-006-358	Sequence 358, App	491	36	49.3	225	14	US-10-207-916-358	Sequence 358, App
419	36	49.3	225	14	US-10-192-008-358	Sequence 358, App	492	36	49.3	225	14	US-10-207-917-358	Sequence 358, App
420	36	49.3	225	14	US-10-192-009-358	Sequence 358, App	493	36	49.3	225	14	US-10-207-918-358	Sequence 358, App
421	36	49.3	225	14	US-10-192-012-358	Sequence 358, App	494	36	49.3	225	14	US-10-207-919-358	Sequence 358, App
422	36	49.3	225	14	US-10-192-014-358	Sequence 358, App	495	36	49.3	225	14	US-10-207-920-358	Sequence 358, App
423	36	49.3	225	14	US-10-192-016-358	Sequence 358, App	496	36	49.3	225	14	US-10-207-925-358	Sequence 358, App
424	36	49.3	225	14	US-10-194-364-358	Sequence 358, App	497	36	49.3	225	14	US-10-208-021-358	Sequence 358, App
425	36	49.3	225	14	US-10-194-364-358	Sequence 358, App	498	36	49.3	225	14	US-10-208-022-358	Sequence 358, App
426	36	49.3	225	14	US-10-194-395-358	Sequence 358, App	499	36	49.3	225	14	US-10-208-023-358	Sequence 358, App
427	36	49.3	225	14	US-10-194-424-358	Sequence 358, App	500	36	49.3	225	14	US-10-208-026-358	Sequence 358, App
428	36	49.3	225	14	US-10-194-458-358	Sequence 358, App							
429	36	49.3	225	14	US-10-194-459-358	Sequence 358, App							
430	36	49.3	225	14	US-10-194-488-358	Sequence 358, App							
431	36	49.3	225	14	US-10-195-886-358	Sequence 358, App							
432	36	49.3	225	14	US-10-195-891-358	Sequence 358, App							
433	36	49.3	225	14	US-10-196-746-358	Sequence 358, App							
434	36	49.3	225	14	US-10-196-752-358	Sequence 358, App							
435	36	49.3	225	14	US-10-196-753-358	Sequence 358, App							
436	36	49.3	225	14	US-10-196-761-358	Sequence 358, App							
437	36	49.3	225	14	US-10-197-692-358	Sequence 358, App							
438	36	49.3	225	14	US-10-197-693-358	Sequence 358, App							
439	36	49.3	225	14	US-10-197-696-358	Sequence 358, App							
440	36	49.3	225	14	US-10-197-698-358	Sequence 358, App							
441	36	49.3	225	14	US-10-197-703-358	Sequence 358, App							
442	36	49.3	225	14	US-10-197-711-358	Sequence 358, App							
443	36	49.3	225	14	US-10-198-757-358	Sequence 358, App							
444	36	49.3	225	14	US-10-198-761-358	Sequence 358, App							
445	36	49.3	225	14	US-10-198-762-358	Sequence 358, App							
446	36	49.3	225	14	US-10-198-763-358	Sequence 358, App							
447	36	49.3	225	14	US-10-198-767-358	Sequence 358, App							
448	36	49.3	225	14	US-10-199-301-358	Sequence 358, App							
449	36	49.3	225	14	US-10-199-307-358	Sequence 358, App							

ALIGNMENTS

RESULT 1

US-10-812-238A-23
; Sequence 23, Application US/10812238A
; Publication No. US20050002904A1
; GENERAL INFORMATION:
; APPLICANT: Wary, Kishore, K.
; APPLICANT: Huntsoe, Joseph O.
; TITLE OF INVENTION: Use of Vascular Endothelial Growth Factor
; TITLE OF INVENTION: and Type I Collagen Inducible Protein (VCIP)
; FILE REFERENCE: D6563
; CURRENT FILING DATE: 2004-03-29
; CURRENT APPLICATION NUMBER: US/10/812,238A
; PRIOR FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: US 60/458,164
; NUMBER OF SEQ ID NOS: 36
; SEQ ID NO 23
; LENGTH: 13

; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: a peptide containing RGD sequence
US-10-812-238A-23

Query Match 100.0%; Score 73; DB 17; Length 13;
Best Local Similarity 100.0%; Pred. No. 2.8e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYRCRGDDSKVQE 13
Db 1 NYRCRGDDSKVQE 13

RESULT 2
US-10-812-238A-2
; Sequence 2, Application US/10812238A
; Publication No. US20050002904A1
; GENERAL INFORMATION:
; APPLICANT: Wary, Kishore, K.
; APPLICANT: Humtsoe, Joseph O.
; TITLE OF INVENTION: Uses of Vascular Endothelial Growth Factor
; TITLE OF INVENTION: and Type I Collagen Inducible Protein (VCIP)
; FILE REFERENCE: D6563
; CURRENT APPLICATION NUMBER: US/10/812,238A
; CURRENT FILING DATE: 2004-03-29
; PRIOR APPLICATION NUMBER: US 60/458,164
; PRIOR FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 36
; SEQ ID NO 2
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: peptide used to raise anti-VCIP-RGD antibody
US-10-812-238A-2

Query Match 100.0%; Score 73; DB 17; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYRCRGDDSKVQE 13
Db 6 NYRCRGDDSKVQE 18

RESULT 3
US-10-655-601-2
; Sequence 2, Application US/10655601
; Publication No. US20040137522A1
; GENERAL INFORMATION:
; APPLICANT: Feany, Mel B.
; APPLICANT: Shulman, Joshua M.
; TITLE OF INVENTION: Genes and Proteins Altering Tau-Related Neuropathy
; FILE REFERENCE: 7570/73251
; CURRENT APPLICATION NUMBER: US/10/655,601
; CURRENT FILING DATE: 2003-09-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-655-601-2

Query Match 100.0%; Score 73; DB 16; Length 311;
Best Local Similarity 100.0%; Pred. No. 0.00063;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYRCRGDDSKVQE 13

Db 178 NYRCRGDDSKVQE 190

RESULT 4
US-10-812-238A-13
; Sequence 13, Application US/10812238A
; Publication No. US20050002904A1
; GENERAL INFORMATION:
; APPLICANT: Wary, Kishore, K.
; APPLICANT: Humtsoe, Joseph O.
; TITLE OF INVENTION: Uses of Vascular Endothelial Growth Factor
; TITLE OF INVENTION: and Type I Collagen Inducible Protein (VCIP)
; FILE REFERENCE: D6563
; CURRENT APPLICATION NUMBER: US/10/812,238A
; CURRENT FILING DATE: 2004-03-29
; PRIOR APPLICATION NUMBER: US 60/458,164
; PRIOR FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 36
; SEQ ID NO 13
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: human VCIP
US-10-812-238A-13

Query Match 100.0%; Score 73; DB 17; Length 311;
Best Local Similarity 100.0%; Pred. No. 0.00063;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYRCRGDDSKVQE 13
Db 178 NYRCRGDDSKVQE 190

RESULT 5
US-10-812-238A-20
; Sequence 20, Application US/10812238A
; Publication No. US20050002904A1
; GENERAL INFORMATION:
; APPLICANT: Wary, Kishore, K.
; APPLICANT: Humtsoe, Joseph O.
; TITLE OF INVENTION: Uses of Vascular Endothelial Growth Factor
; TITLE OF INVENTION: and Type I Collagen Inducible Protein (VCIP)
; FILE REFERENCE: D6563
; CURRENT APPLICATION NUMBER: US/10/812,238A
; CURRENT FILING DATE: 2004-03-29
; PRIOR APPLICATION NUMBER: US 60/458,164
; PRIOR FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 36
; SEQ ID NO 20
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: a peptide containing RGD sequence
US-10-812-238A-20

Query Match 80.8%; Score 59; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0046;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NYRCRGDDSK 10
Db 1 NYRCRGDDSK 10

RESULT 6
US-10-812-238A-22
; Sequence 22, Application US/10812238A

; Publication No. US20050002904A1
; GENERAL INFORMATION:
; APPLICANT: Wary, Kishore, K.
; APPLICANT: Humtsoe, Joseph O.
; TITLE OF INVENTION: Uses of Vascular Endothelial Growth Factor
; TITLE OF INVENTION: and Type I Collagen Inducible Protein (VCIP)
; FILE REFERENCE: D6563
; CURRENT APPLICATION NUMBER: US/10/812,238A
; CURRENT FILING DATE: 2004-03-29
; PRIOR APPLICATION NUMBER: US 60/458,164
; PRIOR FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 36
; SEQ ID NO 22
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: a peptide containing a mutated RGD sequence
US-10-812-238A-22

Query Match 75.3%; Score 55; DB 17; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.021; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
: |||||: |||
Db 1 NYRCRGEDSK 10

RESULT 7
US-10-812-238A-21
; Sequence 21, Application US/10812238A
; Publication No. US20050002904A1
; GENERAL INFORMATION:
; APPLICANT: Wary, Kishore, K.
; APPLICANT: Humtsoe, Joseph O.
; TITLE OF INVENTION: Uses of Vascular Endothelial Growth Factor
; TITLE OF INVENTION: and Type I Collagen Inducible Protein (VCIP)
; FILE REFERENCE: D6563
; CURRENT APPLICATION NUMBER: US/10/812,238A
; CURRENT FILING DATE: 2004-03-29
; PRIOR APPLICATION NUMBER: US 60/458,164
; PRIOR FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 36
; SEQ ID NO 21
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: a peptide containing a mutated RGD sequence
US-10-812-238A-21

Query Match 72.6%; Score 53; DB 17; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.046; Mismatches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
: |||||: |||||
Db 1 NYRCRADDSK 10

RESULT 8
US-10-425-115-262985
; Sequence 262985, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 262985
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_171454C.1.pep
US-10-425-115-262985

Query Match 63.0%; Score 46; DB 16; Length 74;
Best Local Similarity 58.3%; Pred. No. 4.9; Mismatches 4; Mismatches 1; Indels 0; Gaps 0;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 YRCRGDDSKVQE 13
: |||||: |||
Db 42 FLCRGDDSEMQD 53

RESULT 9
US-10-437-963-135805
; Sequence 135805, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 135805
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(208)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_37445C.1.pep
US-10-437-963-135805

Query Match 56.2%; Score 41; DB 16; Length 208;
Best Local Similarity 53.8%; Pred. No. 92; Mismatches 2; Mismatches 4; Indels 0; Gaps 0;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 NYRCRGDDSKVQE 13
: |||||: |||
Db 96 DYSCGDDSGTEE 108

RESULT 10
US-10-291-172-610
; Sequence 610, Application US/10291172
; Publication No. US20030228584A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: No. US20030228584A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-045
; CURRENT APPLICATION NUMBER: US/10/291,172
; CURRENT FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 09/693,267

; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,847
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/596,193
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 610
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-291-172-610

Query Match 56.2%; Score 41; DB 15; Length 237;
Best Local Similarity 70.0%; Pred. No. 1e+02; Mismatches 2; Indels 0; Gaps 0;
Matches 7; Conservative 2

QY 1 NYRCRGDSDK 10
|:|||||:
Db 5 NHRCRGDLTK 14

RESULT 11
US-10-221-278-610
; Sequence 610, Application US/10221278
; Publication No. US20040034208A1
; GENERAL INFORMATION:
; TITLE OF INVENTION: Inc
; FILE REFERENCE: 21272-045
; CURRENT APPLICATION NUMBER: US/10/221,278
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: 09/693,267
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,847
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/596,193
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 610
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-221-278-610

Query Match 56.2%; Score 41; DB 15; Length 237;
Best Local Similarity 70.0%; Pred. No. 1e+02; Mismatches 2; Indels 0; Gaps 0;
Matches 7; Conservative 2

QY 1 NYRCRGDSDK 10
|:|||||:
Db 5 NHRCRGDLTK 14

RESULT 12
US-10-210-281-12
; Sequence 12, Application US/10210281
; Publication No. US20040030096A1
; GENERAL INFORMATION:
; APPLICANT: Gorman, Linda
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Edinger, Shlomit R.

; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Guo, Xiaojia
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Zhong, Mei
; APPLICANT: Patturajan, Meera
; APPLICANT: Miller, Charles E.
; APPLICANT: Ji, Weizhen
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Sciore, Paul
; APPLICANT: Stone, David J.
; APPLICANT: Taupier, Raymond J., Jr.
; APPLICANT: Casman, Stacie
; APPLICANT: Rothenberg, Mark B.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Boldog, Ferenc L.
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
; FILE REFERENCE: 21402-416D
; CURRENT APPLICATION NUMBER: US/10/210,281
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: 60/309,501
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/310,291
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/361,775
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 60/310,951
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/361,832
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 60/311,292
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/311,979
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 60/312,203
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: 60/313,201
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/313,702
; PRIOR FILING DATE: 2001-08-20
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 12
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-210-281-12

Query Match 56.2%; Score 41; DB 15; Length 421;
Best Local Similarity 70.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDSDK 10
|:|||||:
Db 197 NHRCRGDLTK 206

RESULT 13
US-10-210-281-10
; Sequence 10, Application US/10210281
; Publication No. US20040030096A1
; GENERAL INFORMATION:
; APPLICANT: Gorman, Linda
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Guo, Xiaojia
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Zhong, Mei
; APPLICANT: Patturajan, Meera
; APPLICANT: Miller, Charles E.

```

; APPLICANT: Ji, Weizhen
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Sciore, Paul
; APPLICANT: Stone, David J.
; APPLICANT: Taupier, Raymond J., Jr.
; APPLICANT: Casman, Stacie
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Boldog, Ferenc L.
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS
; TITLE OF INVENTION: THE SAME
; FILE REFERENCE: 21402-416D
; CURRENT APPLICATION NUMBER: US/10/210,281
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: 60/309,501
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/310,291
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/361,775
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 60/310,951
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/361,832
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 60/311,292
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/311,979
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 60/312,203
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: 60/313,201
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/313,702
; PRIOR FILING DATE: 2001-08-20
; Remaining Prior Application data removed - See File Wrapper or PALM.
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 10
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-210-281-10

Query Match 56.2%; Score 41; DB 15; Length 424;
Best Local Similarity 70.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
Db 198 NHRCRGDLTK 207
:|||||:|

RESULT 14
US-10-157-031-139
; Sequence 139, Application US/10157031
; Publication No. US20030108890A1
; GENERAL INFORMATION:
; APPLICANT: Baranova, A. V.
; APPLICANT: Iankovsky, N. K.
; APPLICANT: Kozlov, A. P.
; APPLICANT: Lobashev, A. V.
; APPLICANT: Krukovskaya, L. L.
; TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequences
; CURRENT APPLICATION NUMBER: US/10/157,031
; CURRENT FILING DATE: 2002-05-30
; NUMBER OF SEQ ID NOS: 415
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 139
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-157-031-139
Query Match 56.2%; Score 41; DB 14; Length 428;
Best Local Similarity 70.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
Db 197 NHRCRGDLTK 206
:|||||:|

RESULT 15
US-10-437-963-135808
; Sequence 135808, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 135808
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(473)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MET4530_37448C.1.pep
US-10-437-963-135808

Query Match 56.2%; Score 41; DB 16; Length 473;
Best Local Similarity 53.8%; Pred. No. 2.1e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 NYRCRGDDSKVQE 13
Db 153 DYSCGDDSGTEE 165
:|||||:|

RESULT 16
US-10-108-260A-2988
; Sequence 2988, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2988
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-2988

Query Match 56.2%; Score 41; DB 15; Length 489;
Best Local Similarity 70.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```



```
QY      1 NVRCRGDSK 10
      |||||||:
Db      258 NHRCRGLTK 267

RESULT 17
US-09-147-947-6
; Sequence 6, Application US/09147947A
; Patent No. US20020160490A1
; GENERAL INFORMATION:
; APPLICANT: TSURUOKA, No. US20020160490A1uo
; APPLICANT: YAMASHIRO, Kyoto
; APPLICANT: YAMAGUCHI, No. US20020160490A1el Serine Protease
; TITLE OF INVENTION: No. US20020160490A1el Serine Protease
; FILE REFERENCE: 001560-349
; CURRENT APPLICATION NUMBER: US/09/147,947A
; CURRENT FILING DATE: 1997-03-24
; EARLIER APPLICATION NUMBER: PCT/JP98/03324
; EARLIER FILING DATE: 1998-07-24
; EARLIER APPLICATION NUMBER: JP 9/213969
; EARLIER FILING DATE: 1997-07-24
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 822
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; OTHER INFORMATION:
US-09-147-947-6

Query Match      56.2%; Score 41; DB 9; Length 822;
Best Local Similarity 54.5%; Pred. No. 3.5e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 NVRCRGDSK 11
      |||||||:
Db      183 NVRCRDEENI 193

RESULT 18
US-10-757-262-34
; Sequence 34, Application US/10757262
; Publication No. US20040197825A1
; GENERAL INFORMATION:
; APPLICANT: Karicheti, Venkateswarlu
; APPLICANT: Eliasof, Scott D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: UROLOGICAL DISORDERS USING 44390, 54181, 211, 5687, 884,
; TITLE OF INVENTION: 1405, 636, 4421, 5410, 30905, 2045, 16405, 18560, 2047,
; TITLE OF INVENTION: 33751, 52872, 14063, 20739, 32544, 43239, 44373, 51164,
; TITLE OF INVENTION: 53010, 16852, 1587, 2207, 22245, 2387, 52908, 69112, 14990,
; TITLE OF INVENTION: 18547, 115, 579, 15985, 15625, 760, 18603, 2395, 2554, 8675,
; TITLE OF INVENTION: 32720, 4809, 14303, 16816, 17827, 32620, 577, 619, 1423,
; TITLE OF INVENTION: 2158, 8263, 15402, 16209, 16386, 21165, 30911, 41897, 1643,
; TITLE OF INVENTION: 2543, 9626, 13231, 32409, 84260, 2882, 8203, 32678 OR
; FILE REFERENCE: MPI03-007PIRNMNM
; CURRENT APPLICATION NUMBER: US/10/757,262
; CURRENT FILING DATE: 2004-01-14
; PRIOR APPLICATION NUMBER: US 60/440,318
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: US 60/444,783
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: US 60/457,901
; PRIOR FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: US 60/468,775
; PRIOR FILING DATE: 2003-05-08
; PRIOR APPLICATION NUMBER: US 60/471,614
; PRIOR FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: US 60/478,742
; PRIOR FILING DATE: 2003-06-16

QY      1 NVRCRGDSK 11
      |||||||:
Db      183 NVRCRDEENI 193

RESULT 19
US-10-723-860-4046
; Sequence 4046, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193 NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8193
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4046
; LENGTH: 875
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-723-860-4046

Query Match      56.2%; Score 41; DB 16; Length 875;
Best Local Similarity 54.5%; Pred. No. 3.8e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 NVRCRGDSK 11
      |||||||:
Db      236 NVRCRGDEENI 246

RESULT 20
US-10-843-299-2
; Sequence 2, Application US/10843299
; Publication No. US20050032694A1
; GENERAL INFORMATION:
; APPLICANT: SONDEREGGER, Peter
; TITLE OF INVENTION: NEUOTRYPsin
; FILE REFERENCE: 030708-035
; CURRENT APPLICATION NUMBER: US/10/843,299
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/403,724
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/IB98/00625
; PRIOR FILING DATE: 1998-04-24
; PRIOR APPLICATION NUMBER: CH 0966/97
; PRIOR FILING DATE: 1997-04-26
```

```
/ NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 875
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-843-299-2

Query Match          56.2%; Score 41; DB 17; Length 875;
Best Local Similarity 54.5%; Pred. No. 3.8e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 NVRCRGDDSKV 11
   |||||:
Db 236 NVRCRGDBENI 246

RESULT 21
US-10-424-599-260014
; Sequence 260014, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 260014
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(102)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_76817C.1.pep
US-10-424-599-260014

Query Match          54.8%; Score 40; DB 15; Length 102;
Best Local Similarity 70.0%; Pred. No. 67;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 YRCRGDDSKV 11
   |||||:
Db 4 YRCRIDDXRV 13

RESULT 22
US-10-424-599-254922
; Sequence 254922, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 254922
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:

; NAME/KEY: unsure
; LOCATION: (1)-(111)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_72217C.1.pep
US-10-424-599-254922

Query Match          54.8%; Score 40; DB 15; Length 111;
Best Local Similarity 60.0%; Pred. No. 72;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 YRCRGDDSKV 11
   |||||:
Db 66 YXCKGNDTKV 75

RESULT 23
US-10-369-493-15738
; Sequence 15738, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 15738
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Xanthomonas campestris
US-10-369-493-15738

Query Match          54.8%; Score 40; DB 15; Length 301;
Best Local Similarity 50.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 4 CRGDSKVQE 13
   |||||:
Db 200 CRGDDARIHD 209

RESULT 24
US-10-369-493-16124
; Sequence 16124, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 16124
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Xanthomonas campestris
US-10-369-493-16124
```

```
Query Match      54.8%; Score 40; DB 15; Length 301;
Best Local Similarity 50.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY      4 CRGDDSKVOE 13
      |||||...:
Db      200 CRGDDARIHD 209

RESULT 25
US-10-369-493-15371
; Sequence 15371, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 15371
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Xanthomonas campestris
US-10-369-493-15371

Query Match      54.8%; Score 40; DB 15; Length 302;
Best Local Similarity 50.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY      4 CRGDDSKVOE 13
      |||||...:
Db      203 CRGDDARIHD 212

RESULT 26
US-10-282-122A-43293
; Sequence 43293, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
```

```
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43293
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-282-122A-43293

Query Match      54.8%; Score 40; DB 15; Length 400;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2 YRCRGDD 8
      :|||||
Db      39 FRCRGDD 45

RESULT 27
US-10-282-122A-59506
; Sequence 59506, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 59506
```

```
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-10-282-122A-59506

Query Match          54.8%; Score 40; DB 15; Length 400;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2 YRCRGDD 8
       :|||||
Db      39 FRCRGDD 45

RESULT 28
US-10-282-122A-73186
; Sequence 73186, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73186
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Salmonella paratyphi A
US-10-282-122A-73186

Query Match          54.8%; Score 40; DB 15; Length 400;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2 YRCRGDD 8
       :|||||
Db      39 FRCRGDD 45

RESULT 29
US-10-282-122A-75975
; Sequence 75975, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75975
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Salmonella typhi
US-10-282-122A-75975

Query Match          54.8%; Score 40; DB 15; Length 400;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2 YRCRGDD 8
       :|||||
Db      39 FRCRGDD 45

RESULT 30
US-10-425-115-215599
; Sequence 215599, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
```

```
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 215599
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(669)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_128224C.1.pep
US-10-425-115-215599

Query Match          54.8%; Score 40; DB 16; Length 669;
Best Local Similarity 66.7%; Pred. No. 4.3e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 YRCRGDSDK 10
   |||||
Db 379 YKXGDDDK 387

RESULT 31
US-10-437-963-157113
; Sequence 157113, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 157113
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(43)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_56716C.1.pep
US-10-437-963-157113

Query Match          53.4%; Score 39; DB 16; Length 43;
Best Local Similarity 66.7%; Pred. No. 42;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RCRGDDSKV 11
   |||||
Db 29 RCRGDDXEL 37

RESULT 32
US-09-764-877-2043
; Sequence 2043, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper

; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: Patentin Ver. 2.0
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (71)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-242-515-2043

Query Match          53.4%; Score 39; DB 9; Length 74;
Best Local Similarity 63.6%; Pred. No. 71;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 RCRGDDSKVQE 13
   |||||
Db 23 RCGGRDSRSQE 33

RESULT 33
US-10-242-515-2043
; Sequence 2043, Application US/10242515
; Publication No. US20040009488A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005C1
; CURRENT APPLICATION NUMBER: US/10/242,515
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/764,877
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2043
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (71)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-242-515-2043

Query Match          53.4%; Score 39; DB 15; Length 74;
Best Local Similarity 63.6%; Pred. No. 71;
```

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 RCRGDDSKVQE 13
|||:|:|:|:
Db 23 RCGRDSRSQE 33

RESULT 34
US-10-437-963-170560
; Sequence 170560, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 170560
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_68875C.1.pap
US-10-437-963-170560

Query Match 53.4%; Score 39; DB 16; Length 77;
Best Local Similarity 85.7%; Pred. No. 74;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDD 8
:|||||
Db 34 HRCRGDD 40

RESULT 35
US-10-898-775-28
; Sequence 28, Application US/10898775
; Publication No. US20050142647A1
; GENERAL INFORMATION:
; APPLICANT: WASSSELL, PAUL
; APPLICANT: SOE, JORN BORCH
; APPLICANT: MIKKELSON, JORN DALGAARD
; APPLICANT: KRISTENSEN, ANNA CECILIE JENTOF
; TITLE OF INVENTION: METHOD
; FILE REFERENCE: 674509-2058
; CURRENT APPLICATION NUMBER: US/10/898,775
; CURRENT FILING DATE: 2004-07-26
; PRIOR FILING DATE: 2004-07-16
; PRIOR APPLICATION NUMBER: GB 0416023.0
; PRIOR FILING DATE: 2004-01-15
; PRIOR APPLICATION NUMBER: GB 0330016.7
; PRIOR FILING DATE: 2003-12-24
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 28
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Streptomyces coelicolor
US-10-898-775-28

Query Match 53.4%; Score 39; DB 18; Length 305;
Best Local Similarity 63.6%; Pred. No. 2.9e+02;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 RCRGDDSKVQE 13
|||:|:|:|:
Db 258 RCRSDDGAVHE 268

RESULT 36
US-10-417-700A-21
; Sequence 21, Application US/10417700A
; Publication No. US20040033581A1
; GENERAL INFORMATION:
; APPLICANT: ECOPIA BIOSCIENCES INC.
; APPLICANT: ZAZOPOULOS, Emmanuel
; APPLICANT: STAFFA, Alfredo
; APPLICANT: FARNET, Chris
; TITLE OF INVENTION: Specialized dual condensation/epimerization domain in non-ribosome
; TITLE OF INVENTION: Peptide synthetase systems
; FILE REFERENCE: 3002-14US
; CURRENT APPLICATION NUMBER: US/10/417,700A
; CURRENT FILING DATE: 2003-04-17
; NUMBER OF SEQ ID NOS: 139
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; TYPE: PRT
; LENGTH: 461
; ORGANISM: Pseudomonas syringae pv. syringae strain B301D
US-10-417-700A-21

Query Match 53.4%; Score 39; DB 15; Length 461;
Best Local Similarity 50.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDDSKVQ 12
|||:|:|:|:
Db 341 NYRHOGEDNRLQ 352

RESULT 37
US-10-369-493-22730
; Sequence 22730, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 22730
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(690)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-22730

Query Match 53.4%; Score 39; DB 15; Length 690;
Best Local Similarity 80.0%; Pred. No. 6.4e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 RCRGDDSKVQ 12
|||:|:|:|:
Db 561 RVRGDDSDVQ 570

```
RESULT 38
US-10-062-831-131
; Sequence 131, Application US/10062831
; Publication No. US20030105297A1
; GENERAL INFORMATION:
; APPLICANT: Steven M. Ruben, et al.
; TITLE OF INVENTION: 32 Human Secreted Proteins
; FILE REFERENCE: P2006P1
; CURRENT APPLICATION NUMBER: US/10/062,831
; CURRENT FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 09/690,454
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: PCT/US98/10868
; PRIOR FILING DATE: May 28, 1998
; PRIOR APPLICATION NUMBER: 60/044,039
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,093
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,190
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/050,935
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,101
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,356
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/056,250
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,296
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,293
; PRIOR FILING DATE: August 29, 1997
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 131
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-062-831-131

Query Match 52.1%; Score 38; DB 14; Length 136;
Best Local Similarity 54.5%; Pred. No. 1.9e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 YRCRGDDSKVQ 12
|||:|:|
Db 61 YRCKGKNVRVQ 71

RESULT 39
US-10-062-599-131
; Sequence 131, Application US/10062599
; Publication No. US20030195346A1
; GENERAL INFORMATION:
; APPLICANT: Steven M. Ruben, et al.
; TITLE OF INVENTION: 32 Human Secreted Proteins
; FILE REFERENCE: P2006P1
; CURRENT APPLICATION NUMBER: US/10/062,599
; CURRENT FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 09/690,454
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: 09/189,144
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 60/044,039
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,093
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,190
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/050,935
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,101
```

```
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,356
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/056,250
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,296
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,293
; PRIOR FILING DATE: August 29, 1997
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 131
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-062-599-131

Query Match 52.1%; Score 38; DB 14; Length 136;
Best Local Similarity 54.5%; Pred. No. 1.9e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 YRCRGDDSKVQ 12
|||:|:|
Db 61 YRCKGKNVRVQ 71

RESULT 40
US-10-425-115-269918
; Sequence 269918, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 269918
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(151)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MPT4577_177763C.1.pap
US-10-425-115-269918

Query Match 52.1%; Score 38; DB 16; Length 151;
Best Local Similarity 54.5%; Pred. No. 2.1e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYRCRGDDSKV 11
|:|:|:|
Db 132 NHRCKGRDIKL 142

RESULT 41
US-10-282-122A-54254
; Sequence 54254, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
```

```
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54254
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
US-10-282-122A-54254

Query Match          52.1%; Score 38; DB 15; Length 172;
Best Local Similarity 77.8%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 RCRGDDSKV 11
Db 19 RCVGDDGV 27

RESULT 42
US-10-087-192-1548
; Sequence 1548, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 52945200122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1548
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(206)
```

```
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-087-192-1548

Query Match          52.1%; Score 38; DB 13; Length 206;
Best Local Similarity 54.5%; Pred. No. 2.9e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 YRCRGDDSKVQ 12
Db 103 YRCRGKVRVQ 113

RESULT 43
US-10-322-281-322
; Sequence 322, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 322
; LENGTH: 537
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-322-281-322

Query Match          52.1%; Score 38; DB 16; Length 537;
Best Local Similarity 54.5%; Pred. No. 7.4e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 YRCRGDDSKVQ 12
Db 99 YRCRGKVRVQ 109

RESULT 44
US-10-264-049-2872
; Sequence 2872, Application US/10264049
; Publication No. US2004000579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PAL3PI
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2872
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-049-2872

Query Match          52.1%; Score 38; DB 15; Length 545;
Best Local Similarity 54.5%; Pred. No. 7.5e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 YRCRGDDSKVQ 12
Db 103 YRCRGKVRVQ 113

RESULT 45
US-10-264-049-3870
```


; Sequence 3870, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:

; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 3870

; LENGTH: 77

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: MISC_FEATURE

; LOCATION: (1)

; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-049-3870

Query Match 50.7%; Score 37; DB 15; Length 77;
Best Local Similarity 60.8%; Pred. No. 1.6e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYRCRGDGSK 10

Db 41 NPFCKGPDSK 50

RESULT 46

US-09-764-855-84

; Sequence 84, Application US/09764855

; Patent No. US20020119919A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PA110

; CURRENT APPLICATION NUMBER: US/09/764,855

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - consult PALM or file wrapper

; NUMBER OF SEQ ID NOS: 334

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 84

; LENGTH: 113

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-764-855-84

Query Match 50.7%; Score 37; DB 9; Length 113;
Best Local Similarity 62.5%; Pred. No. 2.3e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDD 8

Db 10 NYQCEGED 17

RESULT 47

US-10-072-349-84

; Sequence 84, Application US/10072349

; Publication No. US20030054420A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PA110C1

; CURRENT APPLICATION NUMBER: US/10/072,349

; CURRENT FILING DATE: 2002-02-11

; Prior Application removed - See file Wrapper or Palm

; NUMBER OF SEQ ID NOS: 334

; SOFTWARE: PatentIn Ver. 3.1

; SEQ ID NO 84

; LENGTH: 113

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-072-349-84

Query Match 50.7%; Score 37; DB 14; Length 113;
Best Local Similarity 62.5%; Pred. No. 2.3e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDD 8

Db 10 NYQCEGED 17

RESULT 48

US-10-424-599-165388

; Sequence 165388, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 165388

; LENGTH: 115

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_120361C.1.pap

US-10-424-599-165388

Query Match 50.7%; Score 37; DB 15; Length 115;
Best Local Similarity 54.5%; Pred. No. 2.4e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 YRCRGDSKVQ 12

Db 24 FRCRATNSKLQ 34

RESULT 49

US-10-425-115-283941

; Sequence 283941, Application US/10425115

; Publication No. US20040214272A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B

; CURRENT APPLICATION NUMBER: US/10/425,115

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 369326

; SEQ ID NO 283941

; LENGTH: 133

; TYPE: PRT

; ORGANISM: Zea mays

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1)..(133)

; OTHER INFORMATION: unsure at all Xaa locations

; FEATURE:

; OTHER INFORMATION: Clone ID: MRT4577_22053C.1.pap

US-10-425-115-283941

Query Match 50.7%; Score 37; DB 16; Length 133;
Best Local Similarity 55.6%; Pred. No. 2.7e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 CRGDDSKVQ 12
| | | | |
Db 48 CRGDDDEKIE 56

RESULT 50

US-10-276-774-1857
; Sequence 1857, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y, Tom et al
; TITLE OF INVENTION: No. US20040053245A1 Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 1857
; LENGTH: 140
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(140)
; OTHER INFORMATION: xaa = any amino acid or nothing
US-10-276-774-1857

Query Match 50.7%; Score 37; DB 15; Length 140;
Best Local Similarity 85.7%; Pred. No. 2.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 CRGDDSK 10
| | | | |
Db 80 CRGEDSK 86

Search completed: September 7, 2005, 20:03:13
Job time : 94.1628 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2005, 19:48:00 ; Search time 66.2791 Seconds
(without alignments)
59.419 Million cell updates/sec

Title: US-10-812-238B-20

Perfect score: 59

Sequence: 1 NTRCRGDDSK 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1774312 seqs, 393823214 residues

Total number of hits satisfying chosen parameters: 1774312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database : Published Applications AA.*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10F_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59	100.0	10	17	US-10-812-238A-20
2	59	100.0	13	17	US-10-812-238A-23
3	59	100.0	20	17	US-10-812-238A-2
4	59	100.0	311	16	US-10-655-601-2
5	59	100.0	311	17	US-10-812-238A-13
6	55	93.2	10	17	US-10-812-238A-22
7	53	89.8	10	17	US-10-812-238A-21
8	41	69.5	237	15	US-10-291-172-610
9	41	69.5	237	15	US-10-221-278-610
10	41	69.5	421	15	US-10-210-281-12
11	41	69.5	424	15	US-10-210-281-10
12	385	59.3	269	17	US-10-838-854-78
13	385	59.3	269	17	US-10-478-647-1
14	385	59.3	269	17	US-10-482-029-54
15	385	59.3	269	17	US-09-925-300-1581
16	385	59.3	276	9	US-10-473-127-1710
17	385	59.3	276	16	US-10-437-963-112971
18	385	59.3	293	15	US-10-369-493-15738
19	385	59.3	301	15	US-10-369-493-16124
20	385	59.3	301	15	US-10-369-493-15371
21	385	59.3	302	15	US-10-437-963-170409
22	385	59.3	320	16	US-10-245-752-84
23	385	59.3	385	14	US-10-245-859-84
24	385	59.3	385	14	US-10-245-103-84
25	385	59.3	385	14	US-10-245-103-84
26	385	59.3	385	14	US-10-245-103-84
27	385	59.3	385	14	US-10-245-103-84
28	385	59.3	385	14	US-10-245-103-84
29	385	59.3	385	14	US-10-245-103-84
30	385	59.3	385	14	US-10-245-103-84
31	385	59.3	385	14	US-10-245-103-84
32	385	59.3	385	14	US-10-245-103-84
33	385	59.3	385	14	US-10-245-103-84
34	385	59.3	385	14	US-10-245-103-84
35	385	59.3	385	14	US-10-245-103-84
36	385	59.3	385	14	US-10-245-103-84
37	385	59.3	385	14	US-10-245-103-84
38	385	59.3	385	14	US-10-245-103-84
39	385	59.3	385	14	US-10-245-103-84
40	385	59.3	385	14	US-10-245-103-84
41	385	59.3	385	14	US-10-245-103-84
42	385	59.3	385	14	US-10-245-103-84
43	385	59.3	385	14	US-10-245-103-84
44	385	59.3	385	14	US-10-245-103-84
45	385	59.3	385	14	US-10-245-103-84
46	385	59.3	385	14	US-10-245-103-84
47	385	59.3	385	14	US-10-245-103-84
48	385	59.3	385	14	US-10-245-103-84
49	385	59.3	385	14	US-10-245-103-84
50	385	59.3	385	14	US-10-245-103-84
51	385	59.3	385	14	US-10-245-103-84
52	385	59.3	385	14	US-10-245-103-84
53	385	59.3	385	14	US-10-245-103-84
54	385	59.3	385	14	US-10-245-103-84
55	385	59.3	385	14	US-10-245-103-84
56	385	59.3	385	14	US-10-245-103-84
57	385	59.3	385	14	US-10-245-103-84
58	385	59.3	385	14	US-10-245-103-84
59	385	59.3	385	14	US-10-245-103-84
60	385	59.3	385	14	US-10-245-103-84
61	385	59.3	385	14	US-10-245-103-84
62	385	59.3	385	14	US-10-245-103-84
63	385	59.3	385	14	US-10-245-103-84
64	385	59.3	385	14	US-10-245-103-84
65	385	59.3	385	14	US-10-245-103-84
66	385	59.3	385	14	US-10-245-103-84
67	385	59.3	385	14	US-10-245-103-84
68	385	59.3	385	14	US-10-245-103-84
69	385	59.3	385	14	US-10-245-103-84
70	385	59.3	385	14	US-10-245-103-84
71	385	59.3	385	14	US-10-245-103-84
72	385	59.3	385	14	US-10-245-103-84
73	385	59.3	385	14	US-10-245-103-84
74	385	59.3	385	14	US-10-245-103-84
75	385	59.3	385	14	US-10-245-103-84
76	385	59.3	385	14	US-10-245-103-84
77	385	59.3	385	14	US-10-245-103-84
78	385	59.3	385	14	US-10-245-103-84
79	385	59.3	385	14	US-10-245-103-84
80	385	59.3	385	14	US-10-245-103-84
81	385	59.3	385	14	US-10-245-103-84
82	385	59.3	385	14	US-10-245-103-84
83	385	59.3	385	14	US-10-245-103-84
84	385	59.3	385	14	US-10-245-103-84
85	385	59.3	385	14	US-10-245-103-84
86	385	59.3	385	14	US-10-245-103-84
87	385	59.3	385	14	US-10-245-103-84
88	385	59.3	385	14	US-10-245-103-84
89	385	59.3	385	14	US-10-245-103-84
90	385	59.3	385	14	US-10-245-103-84
91	385	59.3	385	14	US-10-245-103-84
92	385	59.3	385	14	US-10-245-103-84
93	385	59.3	385	14	US-10-245-103-84
94	385	59.3	385	14	US-10-245-103-84
95	385	59.3	385	14	US-10-245-103-84
96	385	59.3	385	14	US-10-245-103-84
97	385	59.3	385	14	US-10-245-103-84
98	385	59.3	385	14	US-10-245-103-84
99	385	59.3	385	14	US-10-245-103-84
100	385	59.3	385	14	US-10-245-103-84
101	385	59.3	385	14	US-10-245-103-84
102	385	59.3	385	14	US-10-245-103-84
103	385	59.3	385	14	US-10-245-103-84
104	385	59.3	385	14	US-10-245-103-84
105	385	59.3	385	14	US-10-245-103-84
106	385	59.3	385	14	US-10-245-103-84
107	385	59.3	385	14	US-10-245-103-84
108	385	59.3	385	14	US-10-245-103-84
109	385	59.3	385	14	US-10-245-103-84
110	385	59.3	385	14	US-10-245-103-84
111	385	59.3	385	14	US-10-245-103-84
112	385	59.3	385	14	US-10-245-103-84
113	385	59.3	385	14	US-10-245-103-84
114	385	59.3	385	14	US-10-245-103-84
115	385	59.3	385	14	US-10-245-103-84
116	385	59.3	385	14	US-10-245-103-84
117	385	59.3	385	14	US-10-245-103-84
118	385	59.3	385	14	US-10-245-103-84
119	385	59.3	385	14	US-10-245-103-84
120	385	59.3	385	14	US-10-245-103-84
121	385	59.3	385	14	US-10-245-103-84
122	385	59.3	385	14	US-10-245-103-84
123	385	59.3	385	14	US-10-245-103-84
124	385	59.3	385	14	US-10-245-103-84
125	385	59.3	385	14	US-10-245-103-84
126	385	59.3	385	14	US-10-245-103-84
127	385	59.3	385	14	US-10-245-103-84
128	385	59.3	385	14	US-10-245-103-84
129	385	59.3	385	14	US-10-245-103-84
130	385	59.3	385	14	US-10-245-103-84
131	385	59.3	385	14	US-10-245-103-84
132	385	59.3	385	14	US-10-245-103-84
133	385	59.3	385	14	US-10-245-103-84
134	385	59.3	385	14	US-10-245-103-84
135	385	59.3	385	14	US-10-245-103-84
136	385	59.3	385	14	US-10-245-103-84
137	385	59.3	385	14	US-10-245-103-84
138	385	59.3	385	14	US-10-245-103-84
139	385	59.3	385	14	US-10-245-103-84
140	385	59.3	385	14	US-10-245-103-84
141	385	59.3	385	14	US-10-245-103-84
142	385	59.3	385	14	US-10-245-103-84
143	385	59.3	385	14	US-10-245-103-84
144	385	59.3	385	14	US-10-245-103-84
145	385	59.3	385	14	US-10-245-103-84
146	385	59.3	385	14	US-10-245-103-84
147	385	59.3	385	14	US-10-245-103-84
148	385	59.3	385	14	US-10-245-103-84
149	385	59.3	385	14	US-10-245-103-84
150	385	59.3	385	14	US-10-245-103-84
151	385	59.3	385	14	US-10-245-103-84
152	385	59.3	385	14	US-10-245-103-84
153	385	59.3	385	14	US-10-245-103-84
154	385	59.3	385	14	US-10-245-103-84
155	385	59.3	385	14	US-10-245-103-84
156	385	59.3	385	14	US-10-245-103-84
157	385	59.3	385	14	US-10-245-103-84
158	385	59.3	385	14	US-10-245-103-84
159	385	59.3	385	14	US-10-245-103-84
160	385	59.3	385	14	US-10-245-103-84
161	385	59.3	385	14	US-10-245-103-84</

85	35	59.3	385	14	US-10-245-107-84	Sequence 84, Appl	158	35	59.3	385	14	US-10-245-479-84	Sequence 84, Appl
86	35	59.3	385	14	US-10-245-143-84	Sequence 84, Appl	159	35	59.3	385	14	US-10-245-499-84	Sequence 84, Appl
87	35	59.3	385	14	US-10-245-771-84	Sequence 84, Appl	160	35	59.3	385	14	US-10-245-772-84	Sequence 84, Appl
88	35	59.3	385	14	US-10-245-851-84	Sequence 84, Appl	161	35	59.3	385	14	US-10-245-811-84	Sequence 84, Appl
89	35	59.3	385	14	US-10-245-883-84	Sequence 84, Appl	162	35	59.3	385	14	US-10-245-812-84	Sequence 84, Appl
90	35	59.3	385	14	US-10-237-535-84	Sequence 84, Appl	163	35	59.3	385	14	US-10-245-852-84	Sequence 84, Appl
91	35	59.3	385	14	US-10-238-183-84	Sequence 84, Appl	164	35	59.3	385	14	US-10-245-875-84	Sequence 84, Appl
92	35	59.3	385	14	US-10-238-283-84	Sequence 84, Appl	165	35	59.3	385	14	US-10-245-881-84	Sequence 84, Appl
93	35	59.3	385	14	US-10-238-370-84	Sequence 84, Appl	166	35	59.3	385	14	US-10-245-911-84	Sequence 84, Appl
94	35	59.3	385	14	US-10-245-055-84	Sequence 84, Appl	167	35	59.3	385	14	US-10-245-913-84	Sequence 84, Appl
95	35	59.3	385	14	US-10-245-147-84	Sequence 84, Appl	168	35	59.3	385	14	US-10-246-080-84	Sequence 84, Appl
96	35	59.3	385	14	US-10-245-730-84	Sequence 84, Appl	169	35	59.3	385	14	US-10-246-121-84	Sequence 84, Appl
97	35	59.3	385	14	US-10-245-739-84	Sequence 84, Appl	170	35	59.3	385	14	US-10-246-305-84	Sequence 84, Appl
98	35	59.3	385	14	US-10-246-210-84	Sequence 84, Appl	171	35	59.3	385	14	US-10-246-929-84	Sequence 84, Appl
99	35	59.3	385	14	US-10-239-196-84	Sequence 84, Appl	172	35	59.3	385	14	US-10-247-036-84	Sequence 84, Appl
100	35	59.3	385	14	US-10-243-024-84	Sequence 84, Appl	173	35	59.3	385	14	US-10-243-255-84	Sequence 84, Appl
101	35	59.3	385	14	US-10-243-409-84	Sequence 84, Appl	174	35	59.3	385	14	US-10-245-810-84	Sequence 84, Appl
102	35	59.3	385	14	US-10-245-621-84	Sequence 84, Appl	175	35	59.3	385	14	US-10-245-910-84	Sequence 84, Appl
103	35	59.3	385	14	US-10-245-880-84	Sequence 84, Appl	176	35	59.3	385	14	US-10-246-098-84	Sequence 84, Appl
104	35	59.3	385	14	US-10-245-033-84	Sequence 84, Appl	177	35	59.3	385	14	US-10-237-496-84	Sequence 84, Appl
105	35	59.3	385	14	US-10-243-095-84	Sequence 84, Appl	178	35	59.3	385	14	US-10-242-074-84	Sequence 84, Appl
106	35	59.3	385	14	US-10-245-185-84	Sequence 84, Appl	179	35	59.3	385	14	US-10-242-505-84	Sequence 84, Appl
107	35	59.3	385	14	US-10-245-427-84	Sequence 84, Appl	180	35	59.3	385	14	US-10-242-574-84	Sequence 84, Appl
108	35	59.3	385	14	US-10-245-473-84	Sequence 84, Appl	181	35	59.3	385	14	US-10-243-261-84	Sequence 84, Appl
109	35	59.3	385	14	US-10-245-770-84	Sequence 84, Appl	182	35	59.3	385	14	US-10-243-282-84	Sequence 84, Appl
110	35	59.3	385	14	US-10-245-877-84	Sequence 84, Appl	183	35	59.3	385	14	US-10-243-402-84	Sequence 84, Appl
111	35	59.3	385	14	US-10-246-976-84	Sequence 84, Appl	184	35	59.3	385	14	US-10-243-431-84	Sequence 84, Appl
112	35	59.3	385	14	US-10-243-320-84	Sequence 84, Appl	185	35	59.3	385	14	US-10-245-164-84	Sequence 84, Appl
113	35	59.3	385	14	US-10-242-743-84	Sequence 84, Appl	186	35	59.3	385	14	US-10-244-972-84	Sequence 84, Appl
114	35	59.3	385	14	US-10-242-845-84	Sequence 84, Appl	187	35	59.3	385	14	US-10-197-942-84	Sequence 84, Appl
115	35	59.3	385	14	US-10-237-636-84	Sequence 84, Appl	188	35	59.3	385	14	US-10-238-196-84	Sequence 84, Appl
116	35	59.3	385	14	US-10-238-325-84	Sequence 84, Appl	189	35	59.3	385	14	US-10-245-013-84	Sequence 84, Appl
117	35	59.3	385	14	US-10-238-346-84	Sequence 84, Appl	190	35	59.3	401	14	US-10-156-761-9009	Sequence 9009, Ap
118	35	59.3	385	14	US-10-238-411-84	Sequence 84, Appl	191	35	59.3	422	14	US-10-203-708-33	Sequence 33, Appl
119	35	59.3	385	14	US-10-243-124-84	Sequence 84, Appl	192	35	59.3	427	10	US-09-907-421-2	Sequence 2, Appl
120	35	59.3	385	14	US-10-243-425-84	Sequence 84, Appl	193	35	59.3	427	18	US-10-771-418-2	Sequence 2, Appl
121	35	59.3	385	14	US-10-243-446-84	Sequence 84, Appl	194	35	59.3	432	15	US-10-108-260A-3406	Sequence 3406, Ap
122	35	59.3	385	14	US-10-245-874-84	Sequence 84, Appl	195	35	59.3	439	14	US-10-139-662-8	Sequence 8, Appl
123	35	59.3	385	14	US-10-242-653-84	Sequence 84, Appl	196	35	59.3	439	14	US-10-139-683-8	Sequence 8, Appl
124	35	59.3	385	14	US-10-243-167-84	Sequence 84, Appl	197	35	59.3	439	14	US-10-230-631-12	Sequence 12, Appl
125	35	59.3	385	14	US-10-243-388-84	Sequence 84, Appl	198	35	59.3	439	14	US-10-143-618-8	Sequence 8, Appl
126	35	59.3	385	14	US-10-244-967-84	Sequence 84, Appl	199	35	59.3	439	16	US-10-745-034-8	Sequence 8, Appl
127	35	59.3	385	14	US-10-244-968-84	Sequence 84, Appl	200	35	59.3	439	16	US-10-777-524-12	Sequence 12, Appl
128	35	59.3	385	14	US-10-244-990-84	Sequence 84, Appl	201	35	59.3	439	16	US-10-745-447-8	Sequence 8, Appl
129	35	59.3	385	14	US-10-245-079-84	Sequence 84, Appl	202	35	59.3	439	17	US-10-777-521-12	Sequence 12, Appl
130	35	59.3	385	14	US-10-245-127-84	Sequence 84, Appl	203	35	59.3	453	15	US-10-369-493-6581	Sequence 6581, Ap
131	35	59.3	385	14	US-10-245-207-84	Sequence 84, Appl	204	35	59.3	453	15	US-10-369-493-6582	Sequence 6582, Ap
132	35	59.3	385	14	US-10-245-646-84	Sequence 84, Appl	205	35	59.3	453	15	US-10-369-493-6583	Sequence 6583, Ap
133	35	59.3	385	14	US-10-245-695-84	Sequence 84, Appl	206	35	59.3	540	16	US-10-437-963-111986	Sequence 111986,
134	35	59.3	385	14	US-10-245-699-84	Sequence 84, Appl	207	35	59.3	552	14	US-10-203-708-34	Sequence 34, Appl
135	35	59.3	385	14	US-10-245-737-84	Sequence 84, Appl	208	35	59.3	553	16	US-10-425-115-352266	Sequence 352266,
136	35	59.3	385	14	US-10-245-878-84	Sequence 84, Appl	209	35	59.3	575	15	US-10-080-334-213	Sequence 213, App
137	35	59.3	385	14	US-10-245-890-84	Sequence 84, Appl	210	35	59.3	575	15	US-10-188-248-34	Sequence 34, Appl
138	35	59.3	385	14	US-10-245-899-84	Sequence 84, Appl	211	35	59.3	615	17	US-10-741-600-1522	Sequence 1522, Ap
139	35	59.3	385	14	US-10-245-900-84	Sequence 84, Appl	212	35	59.3	662	17	US-10-741-600-1523	Sequence 1523, Ap
140	35	59.3	385	14	US-10-247-058-84	Sequence 84, Appl	213	35	59.3	662	17	US-10-741-600-1524	Sequence 1524, Ap
141	35	59.3	385	14	US-10-245-454-84	Sequence 84, Appl	214	35	59.3	702	17	US-10-741-600-1526	Sequence 1526, Ap
142	35	59.3	385	14	US-10-237-471-84	Sequence 84, Appl	215	35	59.3	754	17	US-10-741-600-1525	Sequence 1525, Ap
143	35	59.3	385	14	US-10-238-261-84	Sequence 84, Appl	216	35	59.3	1499	17	US-10-732-923-1744	Sequence 1744, Ap
144	35	59.3	385	14	US-10-238-324-84	Sequence 84, Appl	217	34.5	58.5	1081	16	US-10-437-963-178760	Sequence 178760,
145	35	59.3	385	14	US-10-241-860-84	Sequence 84, Appl	218	34	57.6	51	16	US-10-437-963-204484	Sequence 204484,
146	35	59.3	385	14	US-10-242-172-84	Sequence 84, Appl	219	34	57.6	59	16	US-10-425-115-226744	Sequence 226744,
147	35	59.3	385	14	US-10-243-652-84	Sequence 84, Appl	220	34	57.6	61	16	US-10-425-115-198422	Sequence 198422,
148	35	59.3	385	14	US-10-243-990-84	Sequence 84, Appl	221	34	57.6	93	15	US-10-424-599-177520	Sequence 177520,
149	35	59.3	385	14	US-10-243-023-84	Sequence 84, Appl	222	34	57.6	95	16	US-10-425-115-217038	Sequence 217038,
150	35	59.3	385	14	US-10-243-103-84	Sequence 84, Appl	223	34	57.6	106	15	US-10-424-599-265394	Sequence 265394,
151	35	59.3	385	14	US-10-243-276-84	Sequence 84, Appl	224	34	57.6	113	15	US-10-424-599-241435	Sequence 241435,
152	35	59.3	385	14	US-10-243-326-84	Sequence 84, Appl	225	34	57.6	127	10	US-09-764-891-4168	Sequence 4168, Ap
153	35	59.3	385	14	US-10-243-364-84	Sequence 84, Appl	226	34	57.6	135	15	US-10-316-194-17	Sequence 17, Appl
154	35	59.3	385	14	US-10-243-494-84	Sequence 84, Appl	227	34	57.6	137	15	US-10-732-180-17	Sequence 17, Appl
155	35	59.3	385	14	US-10-244-995-84	Sequence 84, Appl	228	34	57.6	137	15	US-10-104-047-3193	Sequence 3193, Ap
156	35	59.3	385	14	US-10-245-230-84	Sequence 84, Appl	229	34	57.6	142	15	US-10-424-599-176914	Sequence 176914,
157	35	59.3	385	14	US-10-245-253-84	Sequence 84, Appl	230	34	57.6	152	15	US-10-424-599-243663	Sequence 243663,

231	34	57.6	158	16	US-10-767-701-57034	Sequence 57034, A	304	33	55.9	293	16	US-10-437-963-103596	Sequence 103596,
232	34	57.6	172	15	US-10-282-122A-54254	Sequence 54254, A	305	33	55.9	344	15	US-10-424-599-200984	Sequence 200984,
233	34	57.6	202	15	US-10-425-114-65359	Sequence 65359, A	306	33	55.9	355	15	US-10-425-114-49444	Sequence 49444, A
234	34	57.6	257	15	US-10-424-599-197883	Sequence 197883, A	307	33	55.9	360	15	US-10-424-599-228246	Sequence 228246,
235	34	57.6	259	15	US-10-424-599-227542	Sequence 227542, A	308	33	55.9	460	9	US-09-864-761-43042	Sequence 43042, A
236	34	57.6	280	13	US-10-115-899-5	Sequence 5, Appli	309	33	55.9	461	15	US-10-417-700A-21	Sequence 21, Appl
237	34	57.6	280	13	US-10-115-899-8	Sequence 8, Appli	310	33	55.9	469	16	US-10-437-963-103600	Sequence 103600,
238	34	57.6	280	15	US-10-203-534-6	Sequence 6, Appli	311	33	55.9	490	15	US-10-080-334-250	Sequence 250, App
239	34	57.6	280	16	US-10-868-270-1	Sequence 1, Appli	312	33	55.9	490	15	US-10-108-260A-4335	Sequence 4335, App
240	34	57.6	280	16	US-10-868-270-6	Sequence 6, Appli	313	33	55.9	490	15	US-10-275-595A-23	Sequence 23, Appl
241	34	57.6	291	16	US-10-424-599-228719	Sequence 228719, A	314	33	55.9	526	16	US-10-437-963-103644	Sequence 103644,
242	34	57.6	297	16	US-10-437-963-110640	Sequence 110640, A	315	33	55.9	629	16	US-10-375-010-24	Sequence 24, Appl
243	34	57.6	321	18	US-10-617-320-5053	Sequence 5053, Ap	316	33	55.9	654	15	US-10-369-493-9157	Sequence 9157, Ap
244	34	57.6	323	15	US-10-369-493-13114	Sequence 13114, A	317	33	55.9	654	15	US-10-369-493-9433	Sequence 9433, Ap
245	34	57.6	323	16	US-10-437-963-133381	Sequence 133381, A	318	33	55.9	654	15	US-10-369-493-15638	Sequence 15638, A
246	34	57.6	362	16	US-10-737-450-58	Sequence 58, Appl	319	33	55.9	654	15	US-10-369-493-16013	Sequence 16013, A
247	34	57.6	393	16	US-10-425-115-330473	Sequence 330473, A	320	33	55.9	654	15	US-10-369-493-16366	Sequence 16366, A
248	34	57.6	427	16	US-10-739-930-6640	Sequence 6640, Ap	321	33	55.9	675	15	US-10-369-493-2810	Sequence 2810, Ap
249	34	57.6	447	13	US-10-047-676A-10	Sequence 10, Appl	322	33	55.9	678	15	US-10-264-049-3063	Sequence 3063, Ap
250	34	57.6	447	16	US-10-790-914-10	Sequence 10, Appl	323	33	55.9	680	15	US-10-369-493-17753	Sequence 17753, A
251	34	57.6	487	16	US-10-437-963-146626	Sequence 146626, A	324	33	55.9	713	16	US-10-437-963-113424	Sequence 113424,
252	34	57.6	490	14	US-10-156-761-10858	Sequence 10858, A	325	33	55.9	730	14	US-10-067-632-58	Sequence 58, Appl
253	34	57.6	497	16	US-10-437-963-111779	Sequence 111779, A	326	33	55.9	839	16	US-10-437-963-129703	Sequence 129703,
254	34	57.6	510	15	US-10-424-599-206341	Sequence 206341, A	327	33	55.9	842	9	US-09-815-242-11950	Sequence 11950, A
255	34	57.6	541	15	US-10-425-114-72778	Sequence 72778, A	328	33	55.9	842	15	US-10-282-122A-66544	Sequence 66544, A
256	34	57.6	646	16	US-10-437-963-164386	Sequence 164386, A	329	33	55.9	897	15	US-10-259-194A-304	Sequence 304, App
257	34	57.6	801	9	US-09-815-242-5518	Sequence 5518, Ap	330	33	55.9	974	16	US-10-437-963-150927	Sequence 150927,
258	34	57.6	801	9	US-10-282-122A-44225	Sequence 44225, A	331	33	55.9	1004	15	US-10-114-270-32	Sequence 32, Appl
259	34	57.6	801	15	US-10-282-122A-44225	Sequence 44225, A	332	33	55.9	1138	14	US-10-128-714-8165	Sequence 8165, Ap
260	34	57.6	808	15	US-10-104-047-2324	Sequence 2324, Ap	333	33	55.9	1231	17	US-10-868-577A-52	Sequence 52, Appl
261	34	57.6	908	16	US-10-739-930-6350	Sequence 6350, Ap	334	33	55.9	1231	17	US-10-952-459-20	Sequence 20, Appl
262	34	57.6	1036	14	US-10-137-129A-4	Sequence 2, Appli	335	33	55.9	1645	14	US-10-156-761-12609	Sequence 12609, A
263	34	57.6	1070	16	US-10-425-115-314249	Sequence 314249, A	336	33	55.9	1756	16	US-10-425-115-300674	Sequence 300674,
264	34	57.6	1074	14	US-10-137-129A-2	Sequence 2, Appli	337	33	55.9	1907	9	US-09-832-292-39	Sequence 39, Appl
265	34	57.6	1365	16	US-10-437-963-191016	Sequence 191016, A	338	33	55.9	3389	15	US-10-016-248-47	Sequence 47, Appl
266	34	57.6	1687	16	US-10-425-115-215507	Sequence 215507, A	339	33	55.9	3389	16	US-10-408-765A-2286	Sequence 2286, Ap
267	33.5	56.8	223	16	US-10-437-963-204891	Sequence 204891, A	340	33	55.9	3508	15	US-10-016-248-46	Sequence 46, Appl
268	33.5	56.8	308	16	US-10-425-115-292396	Sequence 292396, A	341	33	55.9	4753	15	US-10-369-493-5119	Sequence 5119, Ap
269	33	55.9	356	15	US-10-424-599-268486	Sequence 268486, A	342	32.5	55.1	214	16	US-10-425-115-300451	Sequence 300451,
270	33	55.9	57	16	US-10-425-115-315720	Sequence 315720, A	343	32.5	55.1	255	16	US-10-437-963-116185	Sequence 116185,
271	33	55.9	62	15	US-10-425-115-228943	Sequence 228943, A	344	32.5	55.1	369	16	US-10-425-115-300452	Sequence 300452,
272	33	55.9	69	16	US-10-767-701-61513	Sequence 61513, A	345	32	54.2	18	15	US-10-430-685-26	Sequence 26, Appl
273	33	55.9	78	16	US-10-425-115-24411	Sequence 24411, A	346	32	54.2	20	9	US-09-203-658-27	Sequence 27, Appl
274	33	55.9	88	16	US-10-437-963-190890	Sequence 190890, A	347	32	54.2	20	9	US-09-844-353A-27	Sequence 27, Appl
275	33	55.9	96	16	US-10-437-963-201926	Sequence 201926, A	348	32	54.2	20	10	US-09-963-693-27	Sequence 27, Appl
276	33	55.9	97	16	US-10-437-963-194653	Sequence 194653, A	349	32	54.2	35	9	US-09-205-658-30	Sequence 30, Appl
277	33	55.9	99	16	US-10-693-057-421	Sequence 421, App	350	32	54.2	35	9	US-09-844-353A-30	Sequence 30, Appl
278	33	55.9	99	16	US-10-693-056-421	Sequence 421, App	351	32	54.2	35	10	US-09-963-693-30	Sequence 30, Appl
279	33	55.9	99	17	US-10-840-723-421	Sequence 421, App	352	32	54.2	35	16	US-10-693-057-207	Sequence 207, App
280	33	55.9	99	17	US-10-871-602-421	Sequence 421, App	353	32	54.2	35	17	US-10-693-056-207	Sequence 207, App
281	33	55.9	99	18	US-10-971-679-421	Sequence 421, App	354	32	54.2	35	17	US-10-840-723-207	Sequence 207, App
282	33	55.9	102	16	US-10-437-963-195215	Sequence 195215, A	355	32	54.2	35	17	US-10-871-602-207	Sequence 207, App
283	33	55.9	123	15	US-10-424-599-200987	Sequence 200987, A	356	32	54.2	35	18	US-10-971-679-207	Sequence 207, App
284	33	55.9	155	16	US-10-425-115-261512	Sequence 261512, A	357	32	54.2	43	9	US-09-205-658-28	Sequence 28, Appl
285	33	55.9	159	16	US-10-425-115-346856	Sequence 346856, A	358	32	54.2	43	9	US-09-844-353A-28	Sequence 28, Appl
286	33	55.9	184	18	US-10-499-353A-520	Sequence 520, App	359	32	54.2	43	10	US-09-963-693-28	Sequence 28, Appl
287	33	55.9	185	15	US-10-282-122A-63885	Sequence 63885, A	360	32	54.2	46	11	US-09-985-153-127	Sequence 127, App
288	33	55.9	191	15	US-10-114-270-104	Sequence 104, App	361	32	54.2	55	15	US-10-425-114-37221	Sequence 37221, A
289	33	55.9	201	15	US-10-434-817-12	Sequence 12, Appl	362	32	54.2	59	15	US-10-424-599-203266	Sequence 203266, A
290	33	55.9	203	15	US-11-086-903-12	Sequence 12, Appl	363	32	54.2	66	14	US-10-029-386-29512	Sequence 29512, A
291	33	55.9	203	15	US-10-424-599-234290	Sequence 234290, A	364	32	54.2	67	15	US-10-424-599-238548	Sequence 238548,
292	33	55.9	227	15	US-10-424-599-146170	Sequence 146170, A	365	32	54.2	67	16	US-10-437-963-168930	Sequence 168930,
293	33	55.9	236	15	US-10-434-817-7	Sequence 7, Appli	366	32	54.2	68	15	US-10-424-599-201582	Sequence 201582,
294	33	55.9	236	20	US-11-086-903-7	Sequence 7, Appli	367	32	54.2	69	15	US-10-424-599-213731	Sequence 213731,
295	33	55.9	248	15	US-10-369-493-839	Sequence 839, App	368	32	54.2	73	16	US-10-767-701-53936	Sequence 53936, A
296	33	55.9	255	15	US-10-434-817-2	Sequence 2, Appli	369	32	54.2	74	15	US-10-424-599-203294	Sequence 203294,
297	33	55.9	255	20	US-11-086-903-2	Sequence 2, Appli	370	32	54.2	75	15	US-10-424-599-148291	Sequence 148291,
298	33	55.9	259	16	US-10-437-963-116233	Sequence 116233, A	371	32	54.2	81	16	US-10-437-963-113416	Sequence 113416,
299	33	55.9	262	16	US-10-425-115-338892	Sequence 338892, A	372	32	54.2	86	15	US-10-424-599-258342	Sequence 258342,
300	33	55.9	262	16	US-10-437-963-172864	Sequence 172864, A	373	32	54.2	88	15	US-10-425-114-44290	Sequence 44290, A
301	33	55.9	279	14	US-10-156-761-7657	Sequence 7657, Ap	374	32	54.2	94	18	US-10-972-195-49	Sequence 49, Appl
302	33	55.9	282	15	US-10-424-599-263271	Sequence 263271, A	375	32	54.2	96	16	US-10-425-115-266104	Sequence 266104,
303	33	55.9	282	16	US-10-425-115-338896	Sequence 338896, A	376	32	54.2	97	11	US-09-855-604-369	Sequence 369, App

377 32 54.2 97 12 US-09-855-604-369 Sequence 369, App
378 32 54.2 97 16 US-10-693-057-507 Sequence 507, App
379 32 54.2 97 17 US-10-693-056-507 Sequence 507, App
380 32 54.2 97 17 US-10-840-723-507 Sequence 507, App
381 32 54.2 97 17 US-10-871-602-507 Sequence 507, App
382 32 54.2 97 18 US-10-971-679-507 Sequence 507, App
383 32 54.2 98 16 US-10-425-115-291707 Sequence 291707, App
384 32 54.2 101 16 US-10-693-057-511 Sequence 511, App
385 32 54.2 101 17 US-10-693-056-511 Sequence 511, App
386 32 54.2 101 17 US-10-840-723-511 Sequence 511, App
387 32 54.2 101 17 US-10-871-602-511 Sequence 511, App
388 32 54.2 101 18 US-10-971-679-511 Sequence 511, App
389 32 54.2 102 15 US-10-424-599-245232 Sequence 245232, App
390 32 54.2 107 16 US-10-767-701-57415 Sequence 57415, A
391 32 54.2 110 16 US-10-437-963-103496 Sequence 103496, A
392 32 54.2 111 16 US-10-739-930-6039 Sequence 6039, App
393 32 54.2 112 13 US-10-001-879-183 Sequence 183, App
394 32 54.2 115 16 US-10-425-115-330358 Sequence 330358, App
395 32 54.2 119 15 US-10-424-599-156358 Sequence 156358, App
396 32 54.2 119 16 US-10-437-963-196213 Sequence 196213, App
397 32 54.2 119 17 US-10-652-295-1 Sequence 1, Appli
398 32 54.2 120 9 US-09-066-065A-1 Sequence 3, Appli
399 32 54.2 120 9 US-09-066-065A-3 Sequence 4, Appli
400 32 54.2 120 9 US-09-066-065A-4 Sequence 5, Appli
401 32 54.2 120 9 US-09-066-065A-5 Sequence 6, Appli
402 32 54.2 120 9 US-09-066-065A-6 Sequence 7, Appli
403 32 54.2 120 9 US-09-066-065A-7 Sequence 8, Appli
404 32 54.2 120 9 US-09-066-065A-8 Sequence 9, Appli
405 32 54.2 120 9 US-09-066-065A-9 Sequence 10, Appli
406 32 54.2 120 9 US-09-066-065A-10 Sequence 11, Appli
407 32 54.2 120 9 US-09-066-065A-11 Sequence 12, Appli
408 32 54.2 120 9 US-09-066-065A-12 Sequence 13, Appli
409 32 54.2 120 9 US-09-066-065A-13 Sequence 17, Appli
410 32 54.2 120 9 US-09-066-065A-17 Sequence 18, Appli
411 32 54.2 120 9 US-09-066-065A-19 Sequence 19, Appli
412 32 54.2 120 9 US-09-066-065A-20 Sequence 20, Appli
413 32 54.2 120 9 US-09-066-065A-21 Sequence 21, Appli
414 32 54.2 120 14 US-10-150-262-1 Sequence 34, Appli
415 32 54.2 120 14 US-10-155-886-36 Sequence 36, Appli
416 32 54.2 120 14 US-10-155-886-36 Sequence 38, Appli
417 32 54.2 120 14 US-10-155-886-38 Sequence 1, Appli
418 32 54.2 120 16 US-10-655-756-1 Sequence 136, App
419 32 54.2 120 17 US-10-891-658-135 Sequence 9, Appli
420 32 54.2 120 9 US-10-891-658-136 Sequence 2, Appli
421 32 54.2 121 9 US-09-813-398-9 Sequence 3, Appli
422 32 54.2 121 13 US-10-072-681-2 Sequence 3, Appli
423 32 54.2 121 13 US-10-072-681-3 Sequence 2, Appli
424 32 54.2 121 16 US-10-782-261-2 Sequence 3, Appli
425 32 54.2 121 16 US-10-782-261-3 Sequence 9, Appli
426 32 54.2 121 16 US-10-826-324-9 Sequence 30, Appli
427 32 54.2 121 17 US-10-891-658-30 Sequence 235841, App
428 32 54.2 122 15 US-10-424-599-235841 Sequence 382, App
429 32 54.2 128 11 US-09-855-604-382 Sequence 382, App
430 32 54.2 128 12 US-09-855-604-382 Sequence 46724, A
431 32 54.2 132 16 US-10-767-701-46724 Sequence 283941, App
432 32 54.2 132 16 US-10-425-115-283941 Sequence 190733, App
433 32 54.2 136 16 US-10-437-963-150733 Sequence 287113, App
434 32 54.2 141 16 US-10-425-115-288713 Sequence 269282, App
435 32 54.2 147 15 US-10-424-599-269282 Sequence 66236, A
436 32 54.2 149 15 US-10-425-114-66236 Sequence 68259, A
437 32 54.2 151 15 US-10-425-114-68259 Sequence 938, App
438 32 54.2 152 15 US-10-389-566-938 Sequence 172830, App
439 32 54.2 152 16 US-10-437-963-172830 Sequence 2, Appli
440 32 54.2 153 9 US-09-798-338-2 Sequence 32, Appli
441 32 54.2 153 14 US-10-323-046-32 Sequence 160624, App
442 32 54.2 156 16 US-10-437-963-160624 Sequence 4, Appli
443 32 54.2 157 9 US-09-798-338-4 Sequence 34, Appli
444 32 54.2 157 14 US-10-323-046-34 Sequence 69, Appli
445 32 54.2 157 15 US-10-297-229-69 Sequence 162276, App
446 32 54.2 158 15 US-10-424-599-162278 Sequence 1719, App
447 32 54.2 161 16 US-10-723-860-1719 Sequence 48, Appli
448 32 54.2 161 18 US-10-972-195-48

450 32 54.2 161 18 US-10-756-149-5155 Sequence 5155, App
451 32 54.2 163 9 US-09-798-338-6 Sequence 6, Appli
452 32 54.2 163 14 US-10-323-046-24 Sequence 24, Appli
453 32 54.2 167 9 US-09-798-338-8 Sequence 8, Appli
454 32 54.2 167 14 US-10-323-046-30 Sequence 30, Appli
455 32 54.2 170 15 US-10-424-599-191438 Sequence 191438, App
456 32 54.2 171 10 US-09-939-537-24 Sequence 24, Appli
457 32 54.2 171 11 US-09-243-008-24 Sequence 24, Appli
458 32 54.2 171 16 US-10-408-765A-75 Sequence 75, Appli
459 32 54.2 181 15 US-10-424-599-156359 Sequence 156359, App
460 32 54.2 191 16 US-10-425-115-340441 Sequence 340441, App
461 32 54.2 192 15 US-10-424-599-245759 Sequence 245759, App
462 32 54.2 196 16 US-10-408-765A-2170 Sequence 2170, App
463 32 54.2 197 15 US-10-425-114-60425 Sequence 60425, A
464 32 54.2 204 16 US-10-425-115-205209 Sequence 205209, App
465 32 54.2 207 16 US-10-437-963-199332 Sequence 199332, App
466 32 54.2 209 16 US-10-425-115-254822 Sequence 254822, App
467 32 54.2 210 16 US-10-363-829-401 Sequence 401, App
468 32 54.2 214 9 US-09-198-559-1 Sequence 1, Appli
469 32 54.2 214 13 US-10-091-613-1 Sequence 1, Appli
470 32 54.2 214 15 US-10-211-462-55 Sequence 55, Appli
471 32 54.2 214 15 US-10-211-462-95 Sequence 95, Appli
472 32 54.2 215 17 US-10-507-617-59 Sequence 69, Appli
473 32 54.2 223 15 US-10-374-780A-1897 Sequence 1897, App
474 32 54.2 223 18 US-10-938-409-22 Sequence 22, Appli
475 32 54.2 225 9 US-09-764-868-1102 Sequence 1102, App
476 32 54.2 225 10 US-09-862-540-53 Sequence 53, Appli
477 32 54.2 225 15 US-10-276-774-2405 Sequence 2405, App
478 32 54.2 237 15 US-10-403-571-2 Sequence 2, Appli
479 32 54.2 237 16 US-10-437-963-174095 Sequence 174095, App
480 32 54.2 240 16 US-10-425-115-301068 Sequence 301068, App
481 32 54.2 241 8 US-08-450-842-5 Sequence 5, Appli
482 32 54.2 241 10 US-09-788-188-1 Sequence 1, Appli
483 32 54.2 241 10 US-09-788-188-2 Sequence 2, Appli
484 32 54.2 241 10 US-09-882-263-16 Sequence 16, Appli
485 32 54.2 241 14 US-10-150-262-3 Sequence 3, Appli
486 32 54.2 241 14 US-10-155-886-1 Sequence 6, Appli
487 32 54.2 241 14 US-10-155-886-6 Sequence 7, Appli
488 32 54.2 241 14 US-10-155-886-7 Sequence 8, Appli
489 32 54.2 241 14 US-10-155-886-8 Sequence 9, Appli
490 32 54.2 241 15 US-10-371-003-5 Sequence 5, Appli
491 32 54.2 241 16 US-10-748-337-2 Sequence 2, Appli
492 32 54.2 241 16 US-10-723-860-2698 Sequence 2698, App
493 32 54.2 242 13 US-10-072-681-1 Sequence 1, Appli
494 32 54.2 242 16 US-10-782-261-1 Sequence 1, Appli
495 32 54.2 242 16 US-10-425-115-279555 Sequence 279555, App
496 32 54.2 247 16 US-10-437-963-128934 Sequence 128934, App
497 32 54.2 248 16 US-10-425-115-276941 Sequence 276941, App
498 32 54.2 248 16 US-10-425-115-276941 Sequence 72, Appli
499 32 54.2 250 17 US-10-507-617-72 Sequence 75, Appli
500 32 54.2 251 14 US-10-120-414-75

ALIGNMENTS

RESULT 1
US-10-812-238A-20
; Sequence 20, Application US/10812238A
; Publication No. US20050002904A1
; GENERAL INFORMATION:
; APPLICANT: Wary, Kishore, K.
; APPLICANT: Humtsoe, Joseph O.
; TITLE OF INVENTION: Uses of Vascular Endothelial Growth Factor
; TITLE OF INVENTION: and Type I Collagen Inducible Protein (VCIP)
; FILE REFERENCE: D6563
; CURRENT APPLICATION NUMBER: US/10/812,238A
; CURRENT FILING DATE: 2004-03-29
; PRIOR APPLICATION NUMBER: US 60/458,164
; PRIOR FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 36
; SEQ ID NO 20
; LENGTH: 10

```
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: a peptide containing RGD sequence
US-10-812-238A-20

Query Match      100.0%; Score 59; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0035;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
Db 1 NYRCRGDDSK 10

RESULT 2
US-10-812-238A-23
; Sequence 23, Application US/10812238A
; Publication No. US20050002904A1
; GENERAL INFORMATION:
; APPLICANT: Wary, Kishore, K.
; APPLICANT: Humtsoe, Joseph O.
; TITLE OF INVENTION: Uses of Vascular Endothelial Growth Factor
; TITLE OF INVENTION: and Type I Collagen Inducible Protein (VCIP)
; FILE REFERENCE: D6563
; CURRENT APPLICATION NUMBER: US/10/812,238A
; PRIOR FILING DATE: 2004-03-29
; PRIOR FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 36
; SEQ ID NO 23
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: a peptide containing RGD sequence
US-10-812-238A-23

Query Match      100.0%; Score 59; DB 17; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.0045;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
Db 1 NYRCRGDDSK 10

RESULT 3
US-10-812-238A-2
; Sequence 2, Application US/10812238A
; Publication No. US20050002904A1
; GENERAL INFORMATION:
; APPLICANT: Wary, Kishore, K.
; APPLICANT: Humtsoe, Joseph O.
; TITLE OF INVENTION: Uses of Vascular Endothelial Growth Factor
; TITLE OF INVENTION: and Type I Collagen Inducible Protein (VCIP)
; FILE REFERENCE: D6563
; CURRENT APPLICATION NUMBER: US/10/812,238A
; CURRENT FILING DATE: 2004-03-29
; PRIOR APPLICATION NUMBER: US 60/458,164
; PRIOR FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 36
; SEQ ID NO 2
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: peptide used to raise anti-VCIP-RGD antibody
US-10-812-238A-2
```

```
Query Match      100.0%; Score 59; DB 17; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0067;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
Db 6 NYRCRGDDSK 15

RESULT 4
US-10-655-601-2
; Sequence 2, Application US/10655601
; Publication No. US20040137522A1
; GENERAL INFORMATION:
; APPLICANT: Feany, Mel B.
; APPLICANT: Shulman, Joshua M.
; TITLE OF INVENTION: Genes and Proteins Altering Tau-Related Neuropathy
; FILE REFERENCE: 7570/73251
; CURRENT APPLICATION NUMBER: US/10/655,601
; CURRENT FILING DATE: 2003-09-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION:
US-10-655-601-2

Query Match      100.0%; Score 59; DB 16; Length 311;
Best Local Similarity 100.0%; Pred. No. 0.089;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
Db 178 NYRCRGDDSK 187

RESULT 5
US-10-812-238A-13
; Sequence 13, Application US/10812238A
; Publication No. US20050002904A1
; GENERAL INFORMATION:
; APPLICANT: Wary, Kishore, K.
; APPLICANT: Humtsoe, Joseph O.
; TITLE OF INVENTION: Uses of Vascular Endothelial Growth Factor
; TITLE OF INVENTION: and Type I Collagen Inducible Protein (VCIP)
; FILE REFERENCE: D6563
; CURRENT APPLICATION NUMBER: US/10/812,238A
; CURRENT FILING DATE: 2004-03-29
; PRIOR APPLICATION NUMBER: US 60/458,164
; PRIOR FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 36
; SEQ ID NO 13
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: human VCIP
US-10-812-238A-13

Query Match      100.0%; Score 59; DB 17; Length 311;
Best Local Similarity 100.0%; Pred. No. 0.089;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
Db 178 NYRCRGDDSK 187

RESULT 6
US-10-812-238A-22
; Sequence 22, Application US/10812238A
```

; Publication No. US20050002904A1
; GENERAL INFORMATION:
; APPLICANT: Wary, Kishore, K.
; APPLICANT: Humtsoe, Joseph O.
; TITLE OF INVENTION: Uses of Vascular Endothelial Growth Factor
; TITLE OF INVENTION: and Type I Collagen Inducible Protein (VCIP)
; FILE REFERENCE: D6563
; CURRENT APPLICATION NUMBER: US/10/812,238A
; CURRENT FILING DATE: 2004-03-29
; PRIOR APPLICATION NUMBER: US 60/458,164
; PRIOR FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 36
; SEQ ID NO 22
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: a peptide containing a mutated RGD sequence
US-10-812-238A-22

Query Match 93.2%; Score 55; DB 17; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.016; Indels 0; Gaps 0;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
|:|||||:
Db 1 NYRCRGEDSK 10

RESULT 7
US-10-812-238A-21
; Sequence 21, Application US/10812238A
; Publication No. US20050002904A1
; GENERAL INFORMATION:
; APPLICANT: Wary, Kishore, K.
; APPLICANT: Humtsoe, Joseph O.
; TITLE OF INVENTION: Uses of Vascular Endothelial Growth Factor
; TITLE OF INVENTION: and Type I Collagen Inducible Protein (VCIP)
; FILE REFERENCE: D6563
; CURRENT APPLICATION NUMBER: US/10/812,238A
; CURRENT FILING DATE: 2004-03-29
; PRIOR APPLICATION NUMBER: US 60/458,164
; PRIOR FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 36
; SEQ ID NO 21
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: a peptide containing a mutated RGD sequence
US-10-812-238A-21

Query Match 89.8%; Score 53; DB 17; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.035; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
|:|||||:
Db 1 NYRCRADDSK 10

RESULT 8
US-10-291-172-610
; Sequence 610, Application US/10291172
; Publication No. US20030228584A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: No. US20030228584A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-045
; CURRENT APPLICATION NUMBER: US/10/291,172
; CURRENT FILING DATE: 2000-11-08

; PRIOR APPLICATION NUMBER: 09/693,267
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,847
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/596,193
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 610
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-291-172-610

Query Match 69.5%; Score 41; DB 15; Length 237;
Best Local Similarity 70.0%; Pred. No. 69;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
|:|||||:
Db 5 NHRRCRDLTK 14

RESULT 9
US-10-221-278-610
; Sequence 610, Application US/10221278
; Publication No. US20040034208A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: No. US20040034208A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-045
; CURRENT APPLICATION NUMBER: US/10/221,278
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: 09/693,267
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,847
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/596,193
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 610
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-221-278-610

Query Match 69.5%; Score 41; DB 15; Length 237;
Best Local Similarity 70.0%; Pred. No. 69;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
|:|||||:
Db 5 NHRRCRDLTK 14

RESULT 10
US-10-210-281-12
; Sequence 12, Application US/10210281
; Publication No. US20040030096A1
; GENERAL INFORMATION:
; APPLICANT: Gorman, Linda
; APPLICANT: Zethusen, Bryan D.


```

; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Guo, Xiaojia
; APPLICANT: Kekuda, Rameesh
; APPLICANT: Zhong, Mei
; APPLICANT: Patturajan, Meera
; APPLICANT: Miller, Charles E.
; APPLICANT: Ji, Weizhen
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Sciore, David J.
; APPLICANT: Taupier, Raymond J., Jr.
; APPLICANT: Casman, Stacie
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Boldog, Ferenc L.
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS
; FILE REFERENCE: 21402-416D
; CURRENT APPLICATION NUMBER: US/10/210,281
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: 60/309,501
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/310,291
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/361,775
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 60/310,951
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/361,832
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 60/311,292
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/311,979
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 60/312,203
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: 60/313,201
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/313,702
; PRIOR FILING DATE: 2001-08-20
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 12
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-210-281-12

```

```

Query Match      69.5%; Score 41; DB 15; Length 421;
Best Local Similarity 70.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 NYRCRGDDSK 10
      |:|||||:|
DB      197 NHRCRGDLTK 206

```

```

RESULT 11
US-10-210-281-10
; Sequence 10, Application US/10210281
; Publication No. US20040030096A1
; GENERAL INFORMATION:
; APPLICANT: Gorman, Linda
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Guo, Xiaojia
; APPLICANT: Kekuda, Rameesh
; APPLICANT: Zhong, Mei
; APPLICANT: Patturajan, Meera

```

```

; APPLICANT: Miller, Charles E.
; APPLICANT: Ji, Weizhen
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Sciore, Paul
; APPLICANT: Stone, David J.
; APPLICANT: Taupier, Raymond J., Jr.
; APPLICANT: Casman, Stacie
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Boldog, Ferenc L.
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS
; FILE REFERENCE: 21402-416D
; CURRENT APPLICATION NUMBER: US/10/210,281
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: 60/309,501
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/310,291
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/361,775
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 60/310,951
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/361,832
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 60/311,292
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/311,979
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 60/312,203
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: 60/313,201
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/313,702
; PRIOR FILING DATE: 2001-08-20
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 10
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-210-281-10

```

```

Query Match      69.5%; Score 41; DB 15; Length 424;
Best Local Similarity 70.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 NYRCRGDDSK 10
      |:|||||:|
DB      198 NHRCRGDLTK 207

```

```

RESULT 12
US-10-157-031-139
; Sequence 139, Application US/10157031
; Publication No. US20030108890A1
; GENERAL INFORMATION:
; APPLICANT: Baranova, A. V.
; APPLICANT: Yankovsky, N. K.
; APPLICANT: Kozlov, A. P.
; APPLICANT: Lobashev, A. V.
; APPLICANT: Krukovskaya, L. L.
; TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequences
; FILE REFERENCE: 2760-103
; CURRENT APPLICATION NUMBER: US/10/157,031
; CURRENT FILING DATE: 2002-05-30
; NUMBER OF SEQ ID NOS: 415
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 139
; LENGTH: 428
; TYPE: PRT

```


US-10-282-122A-59506

Query Match 67.8%; Score 40; DB 15; Length 400;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDD 8
:|||||
Db 39 FRCRGDD 45

RESULT 16

US-10-282-122A-73186

; Sequence 73186, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari

; APPLICANT: Zyskind, Judith

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert

; APPLICANT: Forsyth, R.

; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 73186

; LENGTH: 400

; TYPE: PRT

; ORGANISM: Salmonella paratyphi A

US-10-282-122A-73186

Query Match 67.8%; Score 40; DB 15; Length 400;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDD 8
:|||||
Db 39 FRCRGDD 45

RESULT 17

US-10-282-122A-75975

; Sequence 75975, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari

; APPLICANT: Zyskind, Judith

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert

; APPLICANT: Forsyth, R.

; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 75975

; LENGTH: 400

; TYPE: PRT

; ORGANISM: Salmonella typhi

US-10-282-122A-75975

Query Match 67.8%; Score 40; DB 15; Length 400;

Best Local Similarity 85.7%; Pred. No. 1.7e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDD 8

:|||||

Db 39 FRCRGDD 45

RESULT 18

US-10-425-115-215599

; Sequence 215599, Application US/10425115

; Publication No. US20040214272A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53222)B

; CURRENT APPLICATION NUMBER: US/10/425,115

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 369326

; SEQ ID NO 215599

; LENGTH: 669

```
/ TYPE: PRT
/ ORGANISM: Zea mays
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (1)..(669)
/ OTHER INFORMATION: unsure at all Xaa locations
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRT4577_128224C.1.pep
US-10-425-115-215599

Query Match          67.8%; Score 40; DB 16; Length 669;
Best Local Similarity 66.7%; Pred. No. 2.7e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 YRCRGDDSK 10
   |:|||||
Db 379 YKXGDDKK 387

RESULT 19
US-10-437-963-170560
/ Sequence 170560, Application US/10437963
/ Publication No. US20040123343A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Wu, Wei
/ APPLICANT: Boukharov, Andrey A.
/ APPLICANT: Barbazuk, Brad
/ APPLICANT: Li, Ping
/ TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53221)B
/ CURRENT APPLICATION NUMBER: US/10/437,963
/ CURRENT FILING DATE: 2003-05-14
/ NUMBER OF SEQ ID NOS: 204966
/ SEQ ID NO 170560
/ LENGTH: 77
/ TYPE: PRT
/ ORGANISM: Oryza sativa
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT4530_68875C.1.pep
US-10-437-963-170560

Query Match          66.1%; Score 39; DB 16; Length 77;
Best Local Similarity 85.7%; Pred. No. 52;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDD 8
   |:|||||
Db 34 HRCRGDD 40

RESULT 20
US-10-437-963-157113
/ Sequence 157113, Application US/10437963
/ Publication No. US20040123343A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Wu, Wei
/ APPLICANT: Boukharov, Andrey A.
/ APPLICANT: Barbazuk, Brad
/ APPLICANT: Li, Ping
/ TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53221)B
/ CURRENT APPLICATION NUMBER: US/10/437,963
/ CURRENT FILING DATE: 2003-05-14
```

```
/ NUMBER OF SEQ ID NOS: 204966
/ SEQ ID NO 157113
/ LENGTH: 43
/ TYPE: PRT
/ ORGANISM: Oryza sativa
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (1)..(43)
/ OTHER INFORMATION: unsure at all Xaa locations
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT4530_56716C.1.pep
US-10-437-963-157113

Query Match          64.4%; Score 38; DB 16; Length 43;
Best Local Similarity 75.0%; Pred. No. 44;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 RCRGDDSK 10
   |:|||||
Db 29 RCRGDDXE 36

RESULT 21
US-10-425-115-262985
/ Sequence 262985, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 262985
/ LENGTH: 74
/ TYPE: PRT
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRT4577_171454C.1.pep
US-10-425-115-262985

Query Match          64.4%; Score 38; DB 16; Length 74;
Best Local Similarity 66.7%; Pred. No. 73;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 YRCRGDDSK 10
   |:|||||
Db 42 FLCRGDDSE 50

RESULT 22
US-09-147-947-6
/ Sequence 6, Application US/09147947A
/ Patent No. US20020160490A1
/ GENERAL INFORMATION:
/ APPLICANT: TSURUOKA, No. US20020160490A1uo
/ APPLICANT: YAMASHIRO, Kyoto
/ APPLICANT: YAMAGUCHI, No. US20020160490A1omi
/ TITLE OF INVENTION: No. US20020160490A1el Serine Protease
/ FILE REFERENCE: 001560-349
/ CURRENT APPLICATION NUMBER: US/09/147,947A
/ CURRENT FILING DATE: 1997-03-24
/ EARLIER APPLICATION NUMBER: PCT/JP98/03324
/ EARLIER FILING DATE: 1998-07-24
/ EARLIER APPLICATION NUMBER: JP 9/213969
/ EARLIER FILING DATE: 1997-07-24
/ NUMBER OF SEQ ID NOS: 6
/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO 6
```

```

; LENGTH: 822
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; OTHER INFORMATION:
US-09-147-947-6

Query Match          64.4%; Score 38; DB 9; Length 822;
Best Local Similarity 75.0%; Pred. No. 7.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVRCRGDD 8
    |||||:
Db 183 NVRCRGDE 190

RESULT 23
US-10-757-262-34
; Sequence 34, Application US/10757262
; Publication No. US20040197825A1
; GENERAL INFORMATION:
; APPLICANT: Karicheti, Venkateswarlu
; APPLICANT: Eliasof, Scott D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: UROLOGICAL DISORDERS USING 44390, 54181, 211, 5687, 884,
; TITLE OF INVENTION: 1405, 636, 4421, 5410, 30905, 2045, 16405, 18560, 2047,
; TITLE OF INVENTION: 33751, 52872, 14063, 20739, 32544, 43239, 44373, 51164,
; TITLE OF INVENTION: 53010, 16852, 1587, 2207, 22245, 2387, 52908, 69112, 14990,
; TITLE OF INVENTION: 18547, 115, 579, 15985, 15625, 760, 18603, 2395, 2554, 8675,
; TITLE OF INVENTION: 32720, 4809, 14303, 15816, 17827, 32620, 577, 619, 1423,
; TITLE OF INVENTION: 2158, 8263, 15402, 16209, 16386, 21165, 30911, 41897, 1643,
; TITLE OF INVENTION: 2543, 9626, 13231, 32409, 84260, 2882, 8203, 32678 OR
; TITLE OF INVENTION: 55053
; FILE REFERENCE: MPI03-007P1RNMNMIM
; CURRENT APPLICATION NUMBER: US/10/757,262
; CURRENT FILING DATE: 2004-01-14
; PRIOR APPLICATION NUMBER: US 60/440,318
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: US 60/444,783
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: US 60/457,901
; PRIOR FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: US 60/468,775
; PRIOR FILING DATE: 2003-05-08
; PRIOR APPLICATION NUMBER: US 60/471,614
; PRIOR FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: US 60/478,742
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/488,529
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/491,156
; PRIOR FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: US 60/499,594
; PRIOR FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 60/506,332
; PRIOR FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 136
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 875
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-757-262-34

Query Match          64.4%; Score 38; DB 16; Length 875;
Best Local Similarity 75.0%; Pred. No. 7.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVRCRGDD 8
    |||||:
Db 236 NVRCRGDE 243

us-10-723-860-4046
; Sequence 4046, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Gineburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4046
; LENGTH: 875
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-723-860-4046

Query Match          64.4%; Score 38; DB 16; Length 875;
Best Local Similarity 75.0%; Pred. No. 7.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVRCRGDD 8
    |||||:
Db 236 NVRCRGDE 243

RESULT 24
US-10-723-860-4046
; Sequence 4046, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Gineburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4046
; LENGTH: 875
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-723-860-4046

Query Match          64.4%; Score 38; DB 16; Length 875;
Best Local Similarity 75.0%; Pred. No. 7.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVRCRGDD 8
    |||||:
Db 236 NVRCRGDE 243

RESULT 25
US-10-843-299-2
; Sequence 2, Application US/10843299
; Publication No. US20050032694A1
; GENERAL INFORMATION:
; APPLICANT: SONDEREGGER, Peter
; TITLE OF INVENTION: NEUROTRYP SIN
; FILE REFERENCE: 030708-035
; CURRENT APPLICATION NUMBER: US/10/843,299
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/403,724
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/IB98/00625
; PRIOR FILING DATE: 1998-04-24
; PRIOR APPLICATION NUMBER: CH 0966/97
; PRIOR FILING DATE: 1997-04-26
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 875
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-843-299-2

Query Match          64.4%; Score 38; DB 17; Length 875;
Best Local Similarity 75.0%; Pred. No. 7.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVRCRGDD 8
    |||||:
Db 236 NVRCRGDE 243

RESULT 26
US-10-264-049-3870
; Sequence 3870, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

```

```
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 3870
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-049-3870

Query Match          62.7%; Score 37; DB 15; Length 77;
Best Local Similarity 60.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
   |:|:|
Db 41 NFFCKGPD SK 50

RESULT 27
US-09-764-855-84
; Sequence 84, Application US/09764855
; Patent No. US20020119919A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA110
; CURRENT APPLICATION NUMBER: US/09/764,855
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 84
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-855-84

Query Match          62.7%; Score 37; DB 9; Length 113;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDD 8
   |:|:|
Db 10 NYQCEGED 17

RESULT 28
US-10-072-349-84
; Sequence 84, Application US/10072349
; Publication No. US20030054420A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA110C1
; CURRENT APPLICATION NUMBER: US/10/072,349
; CURRENT FILING DATE: 2002-02-11
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 84
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```
US-10-072-349-84

Query Match          62.7%; Score 37; DB 14; Length 113;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDD 8
   |:|:|
Db 10 NYQCEGED 17

RESULT 29
US-10-276-774-1857
; Sequence 1857, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NO. US20040053245A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 1857
; LENGTH: 140
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: (1)...(140)
; OTHER INFORMATION: Xaa = any amino acid or nothing
US-10-276-774-1857

Query Match          62.7%; Score 37; DB 15; Length 140;
Best Local Similarity 85.7%; Pred. No. 2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 CRGDDSK 10
   |:|:|
Db 80 CRGEDSK 86

RESULT 30
US-10-425-115-269918
; Sequence 269918, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 269918
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)...(151)
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: MRT4577_177763C.1.pep
US-10-425-115-269918
```

```
Query Match      62.7%; Score 37; DB 16; Length 151;
Best Local Similarity 60.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 NVRCGDDSK 10
Db      132 NHRCKGRDIK 141

RESULT 31
US-10-425-115-296818
; Sequence 296818, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 296818
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_33771C.1.pcp
US-10-425-115-296818

Query Match      62.7%; Score 37; DB 16; Length 184;
Best Local Similarity 85.7%; Pred. No. 2.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      3 RCRGDDS 9
Db      93 RCRGEDS 99

RESULT 32
US-10-425-115-324709
; Sequence 324709, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 324709
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(286)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_59205C.1.pcp
US-10-425-115-324709

Query Match      62.7%; Score 37; DB 16; Length 286;
Best Local Similarity 66.7%; Pred. No. 3.8e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 NVRCGDDS 9
Db      196 NLRCHGEDS 204

RESULT 33
US-10-156-761-9249
; Sequence 9249, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9249
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9249

Query Match      62.7%; Score 37; DB 14; Length 330;
Best Local Similarity 60.0%; Pred. No. 4.4e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 NVRCGDDSK 10
Db      191 SYKCPGDTSK 200

RESULT 34
US-10-739-930-8397
; Sequence 8397, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 8397
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: ZEAMA-23APR03-C6211_2.p
US-10-739-930-8397

Query Match      62.7%; Score 37; DB 16; Length 348;
Best Local Similarity 85.7%; Pred. No. 4.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      3 RCRGDDS 9
Db      253 RCRGEDS 259

RESULT 35
US-10-812-238A-32
; Sequence 32, Application US/10812238A
```

; Publication No. US20050002904A1
; GENERAL INFORMATION:
; APPLICANT: Wary, Kishore, K.
; APPLICANT: Hunte, Joseph O.
; TITLE OF INVENTION: Uses of Vascular Endothelial Growth Factor
; FILE REFERENCE: D6563
; CURRENT APPLICATION NUMBER: US/10/812,238A
; PRIOR FILING DATE: 2004-03-29
; PRIOR FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 36
; SEQ ID NO 32
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: a peptide containing a RGD sequence
US-10-812-238A-32

Query Match 61.0%; Score 36; DB 17; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.6e+06; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0

QY 4 YRCRGGDSK 9
| | | | |
Db 1 CRGDDS 6

RESULT 36
US-10-424-599-260014
; Sequence 260014, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 260014
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(102)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_76817C.1.pap
US-10-424-599-260014

Query Match 61.0%; Score 36; DB 15; Length 102;
Best Local Similarity 66.7%; Pred. No. 2.1e+02; Indels 0; Gaps 0;
Matches 6; Conservative 1; Mismatches 2

QY 2 YRCRGGDSK 10
| | | | |
Db 4 YCRIDDXR 12

RESULT 37
US-10-424-599-254922
; Sequence 254922, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K

; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 254922
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(111)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_72217C.1.pap
US-10-424-599-254922

Query Match 61.0%; Score 36; DB 15; Length 111;
Best Local Similarity 55.6%; Pred. No. 2.3e+02; Indels 0; Gaps 0;
Matches 5; Conservative 3; Mismatches 1

QY 2 YRCRGGDSK 10
| | | | |
Db 66 YXCKGNDTK 74

RESULT 38
US-10-437-963-126010
; Sequence 126010, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 126010
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_28599C.1.pap
US-10-437-963-126010

Query Match 61.0%; Score 36; DB 16; Length 159;
Best Local Similarity 85.7%; Pred. No. 3.2e+02; Indels 0; Gaps 0;
Matches 6; Conservative 1; Mismatches 0

QY 4 CRGDDSK 10
| | | | |
Db 144 CRGDNSK 150

RESULT 39
US-10-437-963-135805
; Sequence 135805, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 135805
LENGTH: 208
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(208)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_37445C.1.pap
US-10-437-963-135805

Query Match 61.0%; Score 36; DB 16; Length 208;
Best Local Similarity 66.7%; Pred. No. 4.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYRCRGDD 9
:|||||
Db 96 DYSCGGDS 104

RESULT 40
US-10-425-115-214055
Sequence 214055, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 214055
LENGTH: 438
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_126814C.1.pap
US-10-425-115-214055

Query Match 61.0%; Score 36; DB 16; Length 438;
Best Local Similarity 100.0%; Pred. No. 8.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CRGDD 9
:|||||
Db 281 CRGDD 286

RESULT 41
US-10-767-701-41046
Sequence 41046, Application US/10767701
Publication No. US20040172684A1
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
FILE REFERENCE: 38-21(53535)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 41046
LENGTH: 445
TYPE: PRT
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: SORBI-28MAY03-C16361_1.pap
US-10-767-701-41046

Query Match 61.0%; Score 36; DB 16; Length 445;
Best Local Similarity 66.7%; Pred. No. 8.6e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDD 9
:|||||
Db 372 NDQCRGDS 380

RESULT 42
US-10-369-493-2996
Sequence 2996, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 2996
LENGTH: 466
TYPE: PRT
ORGANISM: Thermotoga maritima
US-10-369-493-2996

Query Match 61.0%; Score 36; DB 15; Length 466;
Best Local Similarity 71.4%; Pred. No. 9e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDD 8
:|||||
Db 214 FRCRGED 220

RESULT 43
US-10-369-493-2942
Sequence 2942, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374

```
; SEQ ID NO 2942
; LENGTH: 470
; TYPE: PRT
; ORGANISM: Thermotoga maritima
US-10-369-493-2942

Query Match      61.0%; Score 36; DB 15; Length 470;
Best Local Similarity 71.4%; Pred. No. 9e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      2 YRCRGDD 8
DB      214 FRCRGD 220

RESULT 44
US-10-437-963-135808
; Sequence 135808, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 135808
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: (1)..(473)
; LOCATION: (1)..(473)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_37448C.1.pap
US-10-437-963-135808

Query Match      61.0%; Score 36; DB 16; Length 473;
Best Local Similarity 66.7%; Pred. No. 9.1e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 NYRCRGDD 9
DB      153 DYSCGGDD 161

RESULT 45
US-10-108-260A-2858
; Sequence 2858, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2858
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-2858
```

```
Query Match      61.0%; Score 36; DB 15; Length 485;
Best Local Similarity 60.0%; Pred. No. 9.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 NYRCRGDDSK 10
DB      469 NWRCPGTDSE 478

RESULT 46
US-10-427-631-1
; Sequence 1, Application US/10427631
; Publication No. US20030175923A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; TANG, Y. Tom;
; APPLICANT: CORLEY, Neil C.; GUEGLER, Karl J.;
; APPLICANT: BAUGHN, Mariah R.; LAL, Preeti G.;
; APPLICANT: YUE, Henry; HILLMAN, Jennifer L.;
; APPLICANT: AZIMZAI, Yalda
; TITLE OF INVENTION: HUMAN TRANSFERASE PROTEINS
; FILE REFERENCE: PF-0592-1 DIV
; CURRENT APPLICATION NUMBER: US/10/427,631
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: US 09/786,240
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: PCT/US99/20989
; PRIOR FILING DATE: 1999-09-09
; PRIOR APPLICATION NUMBER: US 60/172,220
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: US 60/155,248
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: US 60/133,642
; PRIOR FILING DATE: 1999-05-11
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030175923A1 1632930CD1
US-10-427-631-1

Query Match      61.0%; Score 36; DB 14; Length 498;
Best Local Similarity 60.0%; Pred. No. 9.5e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 NYRCRGDDSK 10
DB      482 NWRCPGTDSE 491

RESULT 47
US-10-425-115-274886
; Sequence 274886, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 274886
; LENGTH: 1075
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
```

```
; NAME/KEY: unsure
; LOCATION: (1)..(1075)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_182286C.1.pep
US-10-425-115-27486

Query Match      61.0%; Score 36; DB 16; Length 1075;
Best Local Similarity 66.7%; Pred. No. 2e+03;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVRCRGDGS 9
Db 751 NGQCRGDES 759

RESULT 48
US-10-424-599-161655
; Sequence 161655, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 161655
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_116993C.1.pep
US-10-424-599-161655

Query Match      59.3%; Score 35; DB 15; Length 48;
Best Local Similarity 60.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 NVRCRGDGS 10
Db 6 NNRKCRGQDLK 15

RESULT 49
US-10-425-115-320012
; Sequence 320012, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 320012
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_54919C.1.pep
US-10-425-115-320012

Query Match      59.3%; Score 35; DB 16; Length 49;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
```

```
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 YRCRGDGS 10
Db 28 YRCSGSQSK 36

RESULT 50
US-10-425-115-249920
; Sequence 249920, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 249920
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_159510C.1.pep
US-10-425-115-249920

Query Match      59.3%; Score 35; DB 16; Length 57;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGD 7
Db 45 YKCRGD 50

Search completed: September 7, 2005, 20:03:05
Job time : 73.2791 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2005, 19:37:54 ; Search time 93.7209 Seconds
(without alignments)
53.647 Million cell updates/sec

Title: US-10-812-238B-23

Perfect score: 73

Sequence: 1 NYRCRGDDSKVQE 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database :

1: Genesecp16Dec04:*
2: Genesecp19808:*
3: Genesecp20008:*
4: Genesecp20018:*
5: Genesecp20028:*
6: Genesecp2003as:*
7: Genesecp2003bs:*
8: Genesecp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	73	100.0	311	2 AAW79286	Human pro
2	73	100.0	311	7 ADE60703	Human pro
3	73	100.0	311	8 ADM72103	Human tau
4	73	100.0	311	8 ADM72103	Human tau
5	73	100.0	311	8 ADM72103	Human tau
6	69	94.5	312	5 ADP24274	PRO polyp
7	69	94.5	312	5 ADP24274	PRO polyp
8	69	94.5	312	5 ADP24274	PRO polyp
9	69	94.5	312	5 ADP24274	PRO polyp
10	41	56.2	237	8 ADJ50947	Human nov
11	41	56.2	237	8 ADJ50947	Human nov
12	41	56.2	237	8 ADJ50947	Human nov
13	41	56.2	237	8 ADJ50947	Human nov
14	41	56.2	237	8 ADJ50947	Human nov
15	41	56.2	237	8 ADJ50947	Human nov
16	41	56.2	237	8 ADJ50947	Human nov
17	41	56.2	237	8 ADJ50947	Human nov
18	41	56.2	237	8 ADJ50947	Human nov
19	41	56.2	237	8 ADJ50947	Human nov
20	41	56.2	237	8 ADJ50947	Human nov
21	41	56.2	237	8 ADJ50947	Human nov
22	41	56.2	237	8 ADJ50947	Human nov
23	41	56.2	237	8 ADJ50947	Human nov
24	41	56.2	237	8 ADJ50947	Human nov
25	40	54.8	282	5 ABB99111	Rat phosph

26	40	54.8	282	7 ADE58904	Rat Prote
27	40	54.8	282	7 ADE58904	Rat Prote
28	40	54.8	282	7 ADE58904	Rat Prote
29	40	54.8	301	8 ADS26705	Bacterial
30	40	54.8	301	8 ADS26705	Bacterial
31	40	54.8	302	8 ADS26338	Bacterial
32	40	54.8	400	6 ABU31582	Protein e
33	40	54.8	400	6 ABU31582	Protein e
34	40	54.8	400	6 ABU48051	Protein e
35	40	54.8	400	6 ABU48051	Protein e
36	40	54.8	424	7 ABO65513	Klebsiell
37	40	54.8	1299	6 ABP72191	Plasmodiu
38	39	53.4	74	4 ABB04096	Human mus
39	39	53.4	74	4 ABB04096	Human mus
40	39	53.4	74	4 ABB04096	Human mus
41	39	53.4	283	7 ADE61256	Human mus
42	39	53.4	305	8 ADR41635	Rat Prote
43	39	53.4	305	8 ADR41635	Rat Prote
44	39	53.4	393	8 ADM48160	Lipid acy
45	39	53.4	461	7 ADG33762	Thermococ
46	39	53.4	690	8 ADS44300	Actinomyc
47	38.5	52.7	1269	7 ADD46237	Bacterial
48	38.5	52.7	1272	7 ADD46237	Bacterial
49	38	52.1	84	4 AAU50236	Human Pro
50	38	52.1	84	4 AAU50236	Human Pro
51	38	52.1	120	3 AAG02326	Propionib
52	38	52.1	136	2 AAW73460	Human sec
53	38	52.1	172	6 ABU26330	Human sec
54	38	52.1	206	7 ABM85711	Protein e
55	38	52.1	543	4 AAB61306	Human tra
56	38	52.1	543	4 AAB61306	Human tra
57	38	52.1	545	5 ABP41740	Human can
58	38	52.1	798	4 ABG24175	Human ova
59	38	52.1	1307	6 ABP72192	Novel hum
60	38	52.1	1365	6 ABP72192	Plasmodiu
61	38	52.1	2647	8 ADP25446	Plasmodiu
62	37	50.7	59	4 AAU49621	Propionib
63	37	50.7	59	4 AAU49621	Propionib
64	37	50.7	59	4 AAU49621	Propionib
65	37	50.7	60	4 AAU58627	Propionib
66	37	50.7	60	4 AAU58627	Propionib
67	37	50.7	77	5 ABP42738	Propionib
68	37	50.7	92	4 AAU61299	Human ova
69	37	50.7	92	4 AAU61299	Human ova
70	37	50.7	113	4 AAM91935	Propionib
71	37	50.7	113	4 AAM91935	Propionib
72	37	50.7	113	4 AAM91935	Propionib
73	37	50.7	140	4 ABB11487	Human col
74	37	50.7	185	6 ABU35961	Human col
75	37	50.7	190	5 AAE18631	Protein e
76	37	50.7	201	2 AAY27061	Equine hi
77	37	50.7	236	2 AAY27060	Recombina
78	37	50.7	255	2 AAY27058	Equine FC
79	37	50.7	272	7 ABO67231	Equine FC
80	37	50.7	275	7 ABO67231	Klebsiell
81	37	50.7	324	5 ABB09053	Klebsiell
82	37	50.7	349	4 AAM42035	Herbicida
83	37	50.7	409	5 ABP73846	Human pol
84	37	50.7	488	7 ABO66647	Candida a
85	37	50.7	801	4 AAU34022	Klebsiell
86	37	50.7	801	4 AAU34022	Staphyloc
87	37	50.7	801	6 ABM70889	Staphyloc
88	37	50.7	801	6 ABM70889	Staphyloc
89	36	49.3	15	8 ABO32557	Protein e
90	36	49.3	15	8 ABO32557	Protein e
91	36	49.3	134	6 ADE58494	Human pol
92	36	49.3	135	8 ADP29987	Human sec
93	36	49.3	162	4 AAG89271	Human sec
94	36	49.3	191	6 ABO32566	Secreted
95	36	49.3	191	6 ABO32566	Secreted
96	36	49.3	191	6 ABO32566	Secreted
97	36	49.3	215	4 AAU06578	Human NOV
98	36	49.3	215	4 AAU06578	Human NOV
99	36	49.3	215	4 AAU06578	Human NOV
100	36	49.3	215	4 AAU06578	Human NOV

99	36	49.3	215	8	ADQ10176	Human pol	172	36	49.3	225	6	ABU81417	Human PRO
100	36	49.3	225	3	AAY51683	Murine cl	173	36	49.3	225	6	ABR60214	Human sec
101	36	49.3	225	3	AAY99433	Human PRO	174	36	49.3	225	6	ABU91019	Human PRO
102	36	49.3	225	4	AAB66182	Protein o	175	36	49.3	225	6	ABR67949	Human sec
103	36	49.3	225	4	AAM78572	Human PRO	176	36	49.3	225	6	ABR65337	Human sec
104	36	49.3	225	4	AAU29202	Human PRO	177	36	49.3	225	6	ABR68559	Human sec
105	36	49.3	225	4	AAH87585	Human PRO	178	36	49.3	225	6	ABR71971	Human sec
106	36	49.3	225	4	AAG89273	Human sec	179	36	49.3	225	6	ABU85451	Human PRO
107	36	49.3	225	4	AAG89285	Human sec	180	36	49.3	225	6	ABU89141	Human sec
108	36	49.3	225	5	ABG95910	Human sec	181	36	49.3	225	6	ABU83221	Human sec
109	36	49.3	225	5	AAE22227	Human cla	182	36	49.3	225	6	ABU95077	Novel hum
110	36	49.3	225	6	ABU58578	Human PRO	183	36	49.3	225	6	ABU90625	Novel hum
111	36	49.3	225	6	ABU88126	Novel hum	184	36	49.3	225	6	ABU84136	Human sec
112	36	49.3	225	6	ABU84441	Human sec	185	36	49.3	225	6	ABU93787	Novel hum
113	36	49.3	225	6	ABR66315	Human sec	186	36	49.3	225	6	ABR65032	Human sec
114	36	49.3	225	6	ABR65705	Human sec	187	36	49.3	225	6	ABO27340	Human sec
115	36	49.3	225	6	ABU99645	Human sec	188	36	49.3	225	6	ABR68864	Human sec
116	36	49.3	225	6	ABU82884	Human PRO	189	36	49.3	225	6	ABO06680	Human sec
117	36	49.3	225	6	ABU90005	Novel hum	190	36	49.3	225	6	ABR99225	Human sec
118	36	49.3	225	6	ABR68254	Human sec	191	36	49.3	225	6	ABU57109	Human PRO
119	36	49.3	225	6	ABU96307	Novel hum	192	36	49.3	225	6	ABU86061	Novel hum
120	36	49.3	225	6	ABU92738	Human sec	193	36	49.3	225	6	ABU82348	Novel hum
121	36	49.3	225	6	ABO08815	Human sec	194	36	49.3	225	6	ABU87359	Human PRO
122	36	49.3	225	6	ABO02867	Human sec	195	36	49.3	225	6	ABU83831	Human sec
123	36	49.3	225	6	ABR75021	Human sec	196	36	49.3	225	6	ABO08205	Human PRO
124	36	49.3	225	6	ABR94783	Human sec	197	36	49.3	225	6	ABU92535	Human sec
125	36	49.3	225	6	ABU85756	Human PRO	198	36	49.3	225	6	ABU81916	Novel hum
126	36	49.3	225	6	ABU98916	Novel hum	199	36	49.3	225	6	ABU66080	Novel hum
127	36	49.3	225	6	ABU98131	Novel hum	200	36	49.3	225	6	ABU81205	Human sec
128	36	49.3	225	6	ABU91837	Novel hum	201	36	49.3	225	6	ABR59909	Human sec
129	36	49.3	225	6	ABU89530	Human PRO	202	36	49.3	225	6	ABU94097	Novel hum
130	36	49.3	225	6	ABU86371	Human sec	203	36	49.3	225	6	ABU99950	Novel hum
131	36	49.3	225	6	ABU67584	Human sec	204	36	49.3	225	6	ABR66620	Human sec
132	36	49.3	225	6	ABU80612	Human PRO	205	36	49.3	225	6	ABR91038	Human sec
133	36	49.3	225	6	ABU90935	Novel hum	206	36	49.3	225	6	ABO53319	Novel hum
134	36	49.3	225	6	ABO33994	Human sec	207	36	49.3	225	6	ABU94465	Human PRO
135	36	49.3	225	6	ABR99530	Human sec	208	36	49.3	225	6	ABU79347	Human PRO
136	36	49.3	225	6	ABR98920	Human sec	209	36	49.3	225	6	ABR66676	Human sec
137	36	49.3	225	6	ABO16443	Human sec	210	36	49.3	225	6	ABU86981	Novel hum
138	36	49.3	225	6	ABR92343	Human sec	211	36	49.3	225	6	ABU94770	Human PRO
139	36	49.3	225	6	ABO18984	Human sec	212	36	49.3	225	6	ABO04697	Human PRO
140	36	49.3	225	6	ABR78405	Human sec	213	36	49.3	225	6	ABR70446	Human sec
141	36	49.3	225	6	ABU72011	Novel hum	214	36	49.3	225	6	ABU98611	Human PRO
142	36	49.3	225	6	ABU85141	Novel hum	215	36	49.3	225	6	ABR66010	Human sec
143	36	49.3	225	6	ABO00280	Novel hum	216	36	49.3	225	6	ABR64727	Human sec
144	36	49.3	225	6	ABO01612	Human sec	217	36	49.3	225	6	ABU79652	Human PRO
145	36	49.3	225	6	ABO02257	Human sec	218	36	49.3	225	6	ABU93043	Human sec
146	36	49.3	225	6	ABU88831	Novel hum	219	36	49.3	225	6	ABU96002	Human PRO
147	36	49.3	225	6	ABU83526	Human sec	220	36	49.3	225	6	ABU91222	Novel hum
148	36	49.3	225	6	ABO06327	Novel hum	221	36	49.3	225	6	ABU90315	Novel hum
149	36	49.3	225	6	ABR59363	Human sec	222	36	49.3	225	6	ABO09730	Human sec
150	36	49.3	225	6	ABO09425	Human sec	223	36	49.3	225	6	ABO11002	Human sec
151	36	49.3	225	6	ABO19289	Novel hum	224	36	49.3	225	6	ABR71056	Human sec
152	36	49.3	225	6	ABO11307	Human sec	225	36	49.3	225	6	ABU98322	Novel hum
153	36	49.3	225	6	ABR66925	Human sec	226	36	49.3	225	6	ABU87664	Human PRO
154	36	49.3	225	6	ABO32562	Secreted	227	36	49.3	225	6	ABU91532	Human PRO
155	36	49.3	225	6	ABO16138	Human sec	228	36	49.3	225	6	ABU89327	Novel hum
156	36	49.3	225	6	ABO13844	Human sec	229	36	49.3	225	6	ABU84746	Human sec
157	36	49.3	225	6	ABU71565	Human sec	230	36	49.3	225	6	ABR69836	Human sec
158	36	49.3	225	6	ABU65747	Human sec	231	36	49.3	225	6	ABU80213	Human PRO
159	36	49.3	225	6	ABO07595	Human PRO	232	36	49.3	225	6	ABU82534	Novel hum
160	36	49.3	225	6	ABO03782	Human sec	233	36	49.3	225	6	ABU93482	Human PRO
161	36	49.3	225	6	ABR67230	Human sec	234	36	49.3	225	6	ABO10035	Human sec
162	36	49.3	225	6	ABO15833	Human sec	235	36	49.3	225	6	ABO09120	Human sec
163	36	49.3	225	6	ABU56114	Human sec	236	36	49.3	225	6	ABU96498	Human PRO
164	36	49.3	225	6	ABU72346	Human PRO	237	36	49.3	225	6	ABU10688	Human sec
165	36	49.3	225	6	ABU65442	Human PRO	238	36	49.3	225	6	ABU72168	Human PRO
166	36	49.3	225	6	ABU95387	Novel hum	239	36	49.3	225	6	ABU95697	Human PRO
167	36	49.3	225	6	ABU71290	Human PRO	240	36	49.3	225	6	ABU96906	Novel hum
168	36	49.3	225	6	ABO07900	Human PRO	241	36	49.3	225	6	ABR70751	Human sec
169	36	49.3	225	6	ABR70141	Human sec	242	36	49.3	225	6	ABO05102	Novel hum
170	36	49.3	225	6	ABR69474	Human sec	243	36	49.3	225	6	ABO08510	Human sec
171	36	49.3	225	6	ABO01615	Human PRO	244	36	49.3	225	6	ABO05717	Human sec

245	36	49.3	225	6	ABR74106	Human sec	318	49.3	225	6	ADA19982	Novel hum
246	36	49.3	225	6	ABR95698	Human sec	319	49.3	225	6	ABO23799	Human sec
247	36	49.3	225	6	ABR80995	Human sec	320	49.3	225	6	ADB17365	Human tra
248	36	49.3	225	6	ABR81300	Human sec	321	49.3	225	6	ABR97285	Human sec
249	36	49.3	225	6	ABM00996	Human sec	322	49.3	225	6	ABR87073	Human sec
250	36	49.3	225	6	ABR88598	Human sec	323	49.3	225	6	ABM11115	Human sec
251	36	49.3	225	6	ABM77419	Human sec	324	49.3	225	6	ABM28259	Human sec
252	36	49.3	225	6	ABO28903	Human sec	325	49.3	225	6	ABO32258	Human sec
253	36	49.3	225	6	ABO31648	Human sec	326	49.3	225	6	ABM15385	Human sec
254	36	49.3	225	6	ABM08065	Human sec	327	49.3	225	6	ABM06540	Human sec
255	36	49.3	225	6	ABO40545	Human sec	328	49.3	225	6	ABM04351	Human sec
256	36	49.3	225	6	ABO35970	Human PRO	329	49.3	225	6	ABM22464	Human sec
257	36	49.3	225	6	ABO44109	Human PRO	330	49.3	225	6	ABM07760	Human sec
258	36	49.3	225	6	ADA78110	Human sec	331	49.3	225	6	ABO40850	Human sec
259	36	49.3	225	6	ABM24904	Human sec	332	49.3	225	6	ABM35497	Human sec
260	36	49.3	225	6	ABO03172	Human sec	333	49.3	225	6	ABM33260	Human sec
261	36	49.3	225	6	ABR90428	Human sec	334	49.3	225	6	ABO52786	Human PRO
262	36	49.3	225	6	ABM17342	Human sec	335	49.3	225	6	ABO50346	Human sec
263	36	49.3	225	6	ABR95088	Human sec	336	49.3	225	6	ABU99340	Human sec
264	36	49.3	225	6	ABR95088	Human sec	337	49.3	225	6	ABO04392	Human sec
265	36	49.3	225	6	ADB17177	Human tra	338	49.3	225	6	ABO06022	Human sec
266	36	49.3	225	6	ABO21631	Human sec	339	49.3	225	6	ABM18562	Human sec
267	36	49.3	225	6	ABR97895	Human sec	340	49.3	225	6	ABR97590	Human sec
268	36	49.3	225	6	ABR87683	Human sec	341	49.3	225	6	ABR80690	Human sec
269	36	49.3	225	6	ABM77724	Human sec	342	49.3	225	6	ABM01301	Human sec
270	36	49.3	225	6	ABM27954	Human sec	343	49.3	225	6	ABR88903	Human sec
271	36	49.3	225	6	ABM06235	Human sec	344	49.3	225	6	ABM13555	Human sec
272	36	49.3	225	6	ABM03741	Human sec	345	49.3	225	6	ABM20939	Human sec
273	36	49.3	225	6	ABM35192	Human sec	346	49.3	225	6	ABO42070	Human sec
274	36	49.3	225	6	ABM26429	Human sec	347	49.3	225	6	ABO42680	Human sec
275	36	49.3	225	6	ABO48211	Human sec	348	49.3	225	6	ABM10200	Human sec
276	36	49.3	225	6	ABR92953	Human sec	349	49.3	225	6	ABO38715	Human sec
277	36	49.3	225	6	ABO24714	Human sec	350	49.3	225	6	ABM32955	Human sec
278	36	49.3	225	6	ABM11725	Human sec	351	49.3	225	6	ABM22769	Human sec
279	36	49.3	225	6	ABM02826	Human sec	352	49.3	225	6	ABM74980	Human sec
280	36	49.3	225	6	ABM16122	Human sec	353	49.3	225	6	ADA79902	Human sec
281	36	49.3	225	6	ABO27683	Human sec	354	49.3	225	6	ABR96370	Human sec
282	36	49.3	225	6	ABM29174	Human sec	355	49.3	225	6	ABM02521	Human sec
283	36	49.3	225	6	ABM07150	Human sec	356	49.3	225	6	ABR86463	Human sec
284	36	49.3	225	6	ABM21244	Human sec	357	49.3	225	6	ABR86768	Human sec
285	36	49.3	225	6	ABM09590	Human sec	358	49.3	225	6	ABM16732	Human sec
286	36	49.3	225	6	ABO41460	Human sec	359	49.3	225	6	ABM29784	Human sec
287	36	49.3	225	6	ABO36275	Human PRO	360	49.3	225	6	ABO29208	Human sec
288	36	49.3	225	6	ABO43804	Human PRO	361	49.3	225	6	ABM23379	Human sec
289	36	49.3	225	6	ABM76504	Human sec	362	49.3	225	6	ABM22159	Human sec
290	36	49.3	225	6	ABM76200	Human sec	363	49.3	225	6	ABM22159	Human sec
291	36	49.3	225	6	ABM25819	Human sec	364	49.3	225	6	ABO37800	Human sec
292	36	49.3	225	6	ABM26124	Human sec	365	49.3	225	6	ABM28564	Human sec
293	36	49.3	225	6	ABO03477	Human sec	366	49.3	225	6	ABM28869	Human sec
294	36	49.3	225	6	ABO02562	Human sec	367	49.3	225	6	ABM66513	Human sec
295	36	49.3	225	6	ABO44298	Human sec	368	49.3	225	6	ABM75895	Human sec
296	36	49.3	225	6	ABR90733	Human sec	369	49.3	225	6	ABM34175	Human sec
297	36	49.3	225	6	ABR73801	Human sec	370	49.3	225	6	ABM34480	Human sec
298	36	49.3	225	6	ABO17053	Human sec	371	49.3	225	6	ABO20411	Human sec
299	36	49.3	225	6	ABR94478	Human sec	372	49.3	225	6	ABO21326	Human sec
300	36	49.3	225	6	ABR75985	Human sec	373	49.3	225	6	ABO22241	Human sec
301	36	49.3	225	6	ABR71361	Human sec	374	49.3	225	6	ADA20154	Novel hum
302	36	49.3	225	6	ABR93258	Human sec	375	49.3	225	6	ABO34226	Human sec
303	36	49.3	225	6	ABR93563	Human sec	376	49.3	225	6	ABR96675	Human sec
304	36	49.3	225	6	ABR87988	Human sec	377	49.3	225	6	ABR85853	Human sec
305	36	49.3	225	6	ABO33675	Novel hum	378	49.3	225	6	ABR99835	Human sec
306	36	49.3	225	6	ABO27988	Human sec	379	49.3	225	6	ABM00691	Human sec
307	36	49.3	225	6	ABO30123	Human sec	380	49.3	225	6	ABM00386	Human sec
308	36	49.3	225	6	ABO33332	Human PRO	381	49.3	225	6	ABO29818	Human sec
309	36	49.3	225	6	ABM05020	Human sec	382	49.3	225	6	ABM23684	Human sec
310	36	49.3	225	6	ABM08980	Human sec	383	49.3	225	6	ABM29479	Human sec
311	36	49.3	225	6	ABO36580	Human sec	384	49.3	225	6	ABO38410	Human sec
312	36	49.3	225	6	ABO35665	Human PRO	385	49.3	225	6	ABO45710	Human PRO
313	36	49.3	225	6	ABO39630	Human sec	386	49.3	225	6	ABM20634	Human sec
314	36	49.3	225	6	ABM10505	Human sec	387	49.3	225	6	ADA81629	Human sec
315	36	49.3	225	6	ABM12030	Human sec	388	49.3	225	6	ABO16748	Human sec
316	36	49.3	225	6	ABO52176	Human PRO	389	49.3	225	6	ABO18374	Human sec
317	36	49.3	225	6	ABO52481	Human PRO	390	49.3	225	6	ABO22801	Human PRO

391	391	36	49.3	225	6	ABO23106	Human	PRO	Abo23106	464	36	49.3	225	6	ADA00451	Human	sec	Ada00451
392	392	36	49.3	225	6	ABR92648	Human	sec	Abr92648	465	36	49.3	225	6	ABM27039	Human	sec	Abm27039
393	393	36	49.3	225	6	ABR81605	Human	sec	Abr81605	466	36	49.3	225	6	ABM03436	Human	sec	Abm03436
394	394	36	49.3	225	6	ABM78029	Human	sec	Abm78029	467	36	49.3	225	6	ABO39935	Human	sec	Abo39935
395	395	36	49.3	225	6	ABR89818	Human	sec	Abr89818	468	36	49.3	225	7	ABO50041	Human	sec	Abo50041
396	396	36	49.3	225	6	ABM26734	Human	sec	Abm26734	469	36	49.3	225	7	ABO50956	Human	sec	Abo50956
397	397	36	49.3	225	6	ABM13860	Human	sec	Abm13860	470	36	49.3	225	7	ABO05412	Human	sec	Abo05412
398	398	36	49.3	225	6	ABO28598	Human	sec	Abo28598	471	36	49.3	225	7	ABR74716	Human	sec	Abr74716
399	399	36	49.3	225	6	ABO30428	Human	sec	Abo30428	472	36	49.3	225	7	ABO44528	Human	sec	Abo44528
400	400	36	49.3	225	6	ABM07455	Human	sec	Abm07455	473	36	49.3	225	7	ABR77195	Human	sec	Abr77195
401	401	36	49.3	225	6	ABM04046	Human	sec	Abm04046	474	36	49.3	225	7	ABM17952	Human	sec	Abm17952
402	402	36	49.3	225	6	ABO37190	Human	sec	Abo37190	475	36	49.3	225	7	ABR96003	Human	sec	Abr96003
403	403	36	49.3	225	6	ABO41765	Human	sec	Abo41765	476	36	49.3	225	7	ABO21936	Human	sec	Abo21936
404	404	36	49.3	225	6	ABO35360	Human	PRO	Abo35360	477	36	49.3	225	7	ABO20106	Human	sec	Abo20106
405	405	36	49.3	225	6	ABM25209	Human	sec	Abm25209	478	36	49.3	225	7	ABO24409	Human	sec	Abo24409
406	406	36	49.3	225	6	ABO47601	Human	sec	Abm47601	479	36	49.3	225	7	ABR86158	Human	sec	Abr86158
407	407	36	49.3	225	6	ABO47906	Human	sec	Abo47906	480	36	49.3	225	7	ABM10810	Human	sec	Abm10810
408	408	36	49.3	225	6	ABO48516	Human	sec	Abo48516	481	36	49.3	225	7	ABM76809	Human	sec	Abm76809
409	409	36	49.3	225	6	ABO51566	Human	PRO	Abo51566	482	36	49.3	225	7	ABR89513	Human	sec	Abr89513
410	410	36	49.3	225	6	ABO51871	Human	PRO	Abo51871	483	36	49.3	225	7	ABM12640	Human	sec	Abm12640
411	411	36	49.3	225	6	ABO50651	Human	sec	Abo50651	484	36	49.3	225	7	ABM05930	Human	sec	Abo05930
412	412	36	49.3	225	6	ABR57975	Human	sec	Abr79775	485	36	49.3	225	7	ABO35055	Human	PRO	Abo35055
413	413	36	49.3	225	6	ABM17037	Human	sec	Abm17037	486	36	49.3	225	7	ABM03131	Human	sec	Abm03131
414	414	36	49.3	225	6	ABO18069	Human	sec	Abo18069	487	36	49.3	225	7	ABM19109	Human	sec	Abm19109
415	415	36	49.3	225	6	ABO21021	Human	sec	Abo21021	488	36	49.3	225	7	ABM19414	Human	sec	Abm19414
416	416	36	49.3	225	6	ABR96980	Human	sec	Abr96980	489	36	49.3	225	7	ABO46625	Human	PRO	Abo46625
417	417	36	49.3	225	6	ABM12335	Human	sec	Abm12335	490	36	49.3	225	7	ABO49126	Human	sec	Abo49126
418	418	36	49.3	225	6	ABM16427	Human	sec	Abm16427	491	36	49.3	225	7	ABR69169	Human	sec	Abr69169
419	419	36	49.3	225	6	ABM24294	Human	sec	Abm24294	492	36	49.3	225	7	ABR89208	Human	sec	Abr89208
420	420	36	49.3	225	6	ABM14775	Human	sec	Abm14775	493	36	49.3	225	7	ABR72581	Human	sec	Abr72581
421	421	36	49.3	225	6	ABM04656	Human	sec	Abm04656	494	36	49.3	225	7	ABR74411	Human	sec	Abr74411
422	422	36	49.3	225	6	ABM06845	Human	sec	Abm06845	495	36	49.3	225	7	ABO18679	Human	sec	Abo18679
423	423	36	49.3	225	6	ABM09285	Human	sec	Abm09285	496	36	49.3	225	7	ABR80385	Human	sec	Abr80385
424	424	36	49.3	225	6	ABO39325	Human	sec	Abo39325	497	36	49.3	225	7	ABM01606	Human	sec	Abm01606
425	425	36	49.3	225	6	ABM75590	Human	sec	Abm75590	498	36	49.3	225	7	ABM02216	Human	sec	Abm02216
426	426	36	49.3	225	6	ABM25514	Human	sec	Abm25514	499	36	49.3	225	7	ABR87378	Human	sec	Abr87378
427	427	36	49.3	225	6	ABM20024	Human	sec	Abm20024	500	36	49.3	225	7	ABM12945	Human	sec	Abm12945
428	428	36	49.3	225	6	ABO46930	Human	PRO	Abo46930									
429	429	36	49.3	225	6	ABO47235	Human	PRO	Abo47235									
430	430	36	49.3	225	6	ADA83427	Human	sec	Ada83427									
431	431	36	49.3	225	6	ABR71666	Human	sec	Abr71666									
432	432	36	49.3	225	6	ABR72276	Human	sec	Abr72276									
433	433	36	49.3	225	6	ABR98615	Human	sec	Abr98615									
434	434	36	49.3	225	6	ABO06985	Human	sec	Abo06985									
435	435	36	49.3	225	6	ABR84938	Human	sec	Abr84938									
436	436	36	49.3	225	6	ABR73496	Human	sec	Abr73496									
437	437	36	49.3	225	6	ABR76590	Human	sec	Abr76590									
438	438	36	49.3	225	6	ABR73191	Human	sec	Abr73191									
439	439	36	49.3	225	6	ABM18257	Human	sec	Abm18257									
440	440	36	49.3	225	6	ABO20716	Human	PRO	Abo20716									
441	441	36	49.3	225	6	ABO25459	Human	PRO	Abo25459									
442	442	36	49.3	225	6	ABO25764	Human	PRO	Abo25764									
443	443	36	49.3	225	6	ABR94173	Human	sec	Abr94173									
444	444	36	49.3	225	6	ABR80080	Human	sec	Abr80080									
445	445	36	49.3	225	6	ABM11420	Human	sec	Abm11420									
446	446	36	49.3	225	6	ABO33027	Human	PRO	Abo33027									
447	447	36	49.3	225	6	ABO30733	Human	sec	Abo30733									
448	448	36	49.3	225	6	ABO31038	Human	sec	Abo31038									
449	449	36	49.3	225	6	ABM27344	Human	sec	Abm27344									
450	450	36	49.3	225	6	ABM30089	Human	sec	Abm30089									
451	451	36	49.3	225	6	ABM05625	Human	sec	Abm05625									
452	452	36	49.3	225	6	ABM15690	Human	sec	Abm15690									
453	453	36	49.3	225	6	ABM08675	Human	sec	Abm08675									
454	454	36	49.3	225	6	ABO42375	Human	sec	Abo42375									
455	455	36	49.3	225	6	ABO38105	Human	sec	Abo38105									
456	456	36	49.3	225	6	ABO46015	Human	PRO	Abo46015									
457	457	36	49.3	225	6	ABM66818	Human	sec	Abm66818									
458	458	36	49.3	225	6	ADB20470	Human	sec	Adb20470									
459	459	36	49.3	225	6	ABM19719	Human	sec	Abm19719									
460	460	36	49.3	225	6	ABO49431	Human	sec	Abo49431									
461	461	36	49.3	225	6	ABO49736	Human	sec	Abo49736									
462	462	36	49.3	225	6	ADA78722	Human	sec	Ada78722									
463	463	36	49.3	225	6	ABR88293	Human	sec	Abr88293									

ALIGNMENTS

RESULT 1

AAW79286

ID AAW79286 standard; protein; 311 AA.

XX

XX

AC AAW79286;

XX

DT 15-FEB-1999 (first entry)

XX

Human phosphatidic acid phosphatase beta.

XX

DE

XX

Phosphatidic acid phosphatase beta; PAP-beta; human; dephosphorylation;

XX

Phosphatidic acid phosphatase beta; PAP-beta; human; dephosphorylation;

XX

tumour suppressor; cancer; gene therapy.

XX

Homo sapiens.

XX

WO9846730-A1.

XX

22-OCT-1998.

XX

16-APR-1998; 98WO-US007928.

XX

17-APR-1997; 97US-00842827.

XX

(CELL-) CELL THERAPEUTICS INC.

XX

Leung DW, Tompkins CK;

XX

WPI; 1998-594568/50.

DR

N-PSDB; AAV69088.

XX

New nucleic acid encoding human phosphatidic acid phosphatases - used to

PT regulate levels of lipid cellular mediators and in gene therapy of e.g.
 XX cancer.
 PS Claim 8; Fig 3A-B; 62pp; English.
 XX This is the amino acid sequence of human mature phosphatidic acid
 CC phosphatase-beta (PAP-beta), an enzyme that catalyses the conversion of
 CC phosphatidic acid into diacylglycerol. 3 Variants of human PAP, i.e. PAP-
 CC alpha 1 and 2 (see AAW79284-85), PAP-beta (see AAW79286) and PAP-gamma
 CC (see AAW79287) have been identified. The invention provides PAP
 CC polynucleotides (see AAV69086-89) and polypeptides, a method of preparing
 CC PAP in a transformed host cell, and a method of using PAP to
 CC dephosphorylate a substrate, especially lysophosphatidic acid, ceramide 1
 CC -phosphate or sphingosine 1-phosphate, particularly for production of
 CC diacylglycerol, but also monoacylglycerol, ceramide and sphingosine. PAP
 CC is able to control the balance of lipid mediators of cellular activation
 CC and signal transduction. Sequences that encode PAP are potentially
 CC useful, in gene therapy, for treatment of cancer (PAP may be tumour
 CC suppressors, PAP-alpha is expressed at lower levels in cancer cells than
 CC in normal cells of same tissue type), inflammatory disease and diabetes-
 CC associated obesity
 XX
 SQ Sequence 311 AA;
 Query Match 100.0%; Score 73; DB 2; Length 311;
 Best Local Similarity 100.0%; Pred. No. 0.00083;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NYRCRGDSDSKVQE 13
 Db 178 NYRCRGDSDSKVQE 190
 |||||
 RESULT 2
 ADE60703
 ID ADE60703 standard; protein; 311 AA.
 XX
 AC ADE60703;
 XX
 XX 29-JAN-2004 (first entry)
 XX Human Protein NP_003704, SEQ ID NO 6615.
 XX
 DE Human; pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; chronic constriction injury; CCI;
 KW spared nerve injury; SNI; Chung.
 XX
 OS Homo sapiens.
 XX
 XX WO2003016475-A2.
 PN 27-FEB-2003.
 XX
 PD 14-AUG-2002; 2002WO-US025765.
 PF
 XX 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX
 XX (GEHO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 XX
 XX Woolf C, D'urso D, Befort K, Costigan M;
 XX WPI; 2003-268312/26.
 DR GENBANK; NP_003704.
 XX
 XX New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 XX
 PS Claim 1; Page; 1017pp; English.
 XX
 XX The invention discloses a composition comprising two or more isolated rat

CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a human protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 311 AA;
 Query Match 100.0%; Score 73; DB 7; Length 311;
 Best Local Similarity 100.0%; Pred. No. 0.00083;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NYRCRGDSDSKVQE 13
 Db 178 NYRCRGDSDSKVQE 190
 |||||
 RESULT 3
 ADM72103
 ID ADM72103 standard; protein; 311 AA.
 XX
 AC ADM72103;
 XX
 XX 03-JUN-2004 (first entry)
 XX Human tau-related polypeptide PPAP2B.
 XX
 DE Tau; PPAP2A; PPAP2B; IHPK1; IHPK3; IHPK2; FLJ20530; D7434014.5; EZF1;
 KW Loc127424; Loc113179; KIAA0826; neuroprotective; neurotropic;
 KW antiparkinsonian; tau-protein kinase; human.
 XX
 OS Homo sapiens.
 XX
 XX WO2004022708-A2.
 PN 18-MAR-2004.
 XX
 PD 04-SEP-2003; 2003WO-US027590.
 PF
 XX 09-SEP-2002; 2002US-0408877P.
 PR
 XX (BGHM) BRIGHAM & WOMENS HOSPITAL INC.
 XX
 XX Feany MB, Shulman JM;
 XX WPI; 2004-248456/23.
 DR
 XX New substantially pure Tau-related polypeptides and polynucleotides,
 PT useful for diagnosing and/or treating neurological conditions with
 PT aberrant expression of the Tau-related polypeptide, such as Alzheimer's
 PT disease.
 XX

PS Claim 1; SEQ ID NO 2; 62pp; English.

XX The invention relates to tau-related polypeptide consisting essentially

CC of an amino acid sequence selected from PPAP2A, PPAP2B, IHPK1, IHPK3,

CC IHPK2, ELJ20530, DJ434014.5, EZF1, LOC127424, LOC113179 and KIAA0826

CC (ADM72102-ADM72112 respectively). The methods and compositions of the

CC present invention are useful for the diagnosis and/or treatment of

CC neurological diseases or conditions associated with aberrant expression

CC or activity of the Tau-related polypeptide, such as Alzheimer's disease

CC and Parkinson's disease. The present sequence represents a human tau-

CC related polypeptide homologue PPAP2B.

XX Sequence 311 AA;

QY Query Match 100.0%; Score 73; DB 8; Length 311;

Best Local Similarity 100.0%; Pred. No. 0.00083;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 NYRCRGDDSKVQE 13

178 NYRCRGDDSKVQE 190

RESULT 4

ADN04415

ID ADN04415 standard; protein; 311 AA.

AC ADN04415;

XX 01-JUL-2004 (first entry)

DE Antipsoriatic protein sequence #401.

KW antipsoriatic; gene therapy; psoriasis; diagnosis.

XX Homo sapiens.

OS WO2004028479-A2.

PN 08-APR-2004.

PD 25-SEP-2003; 2003WO-US030907.

PF 25-SEP-2002; 2002US-0414006P.

XX (GETH) GENENTECH INC.

PA Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;

PI Wu TD;

XX WPI; 2004-305105/28.

DR N-PSDB; ADN04414.

XX New PRO nucleic acid or polypeptide, useful for preparing a

PT pharmaceutical composition for diagnosing or treating psoriasis in a

PT mammal.

XX Claim 9; SEQ ID NO 809; 3069pp; English.

PS The invention relates to novel polynucleotide and polypeptides for

CC treating psoriasis or a sequence having at least 80% identity to the

CC above sequences. The nucleic acid is useful for preparing a composition

CC for diagnosing or treating psoriasis in a mammal. This sequence

CC corresponds to one of the polypeptides of the invention.

XX Sequence 311 AA;

QY Query Match 100.0%; Score 73; DB 8; Length 311;

Best Local Similarity 100.0%; Pred. No. 0.00083;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYRCRGDDSKVQE 13

178 NYRCRGDDSKVQE 190

DB 178 NYRCRGDDSKVQE 190

RESULT 5

ADP24274

ID ADP24274 standard; protein; 311 AA.

XX ADP24274;

AC 18-NOV-2004 (first entry)

DT PRO polypeptide SEQ ID NO:1452.

XX

DE PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;

KW osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;

KW antiasthmatic; hepatotropic; respiratory; gene therapy; immune system.

XX Unidentified.

OS WO2004041170-A2.

PN 21-MAY-2004.

PD 30-OCT-2003; 2003WO-US034312.

PF 01-NOV-2002; 2002US-0423394P.

PR (GETH) GENENTECH INC.

PA Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;

PI Wu TD;

XX WPI; 2004-419628/39.

DR N-PSDB; ADP24273.

XX New PRO polypeptides and polynucleotides, useful for treating e.g.

PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated

PT renal disease, or demyelinating diseases of the central or peripheral

PT nervous system.

XX Claim 7; SEQ ID NO 1452; 2940pp; English.

PS The invention relates to a novel isolated nucleic acid and the PRO

CC polypeptide encoded by it. A protein of the invention has

CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,

CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,

CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide

CC of the invention may have a use in gene therapy. The PRO polypeptide, its

CC agonist, antagonist, or antibody that specifically binds to the

CC polypeptide is useful for treating an immune related disorder such as

CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,

CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an

CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic

CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune

CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal

CC disease, a demyelinating disease of the central or peripheral nervous

CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,

CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary

CC disease, infectious or autoimmune chronic active hepatitis, primary

CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,

CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's

CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic

CC disease, asthma, allergic rhinitis, atopic dermatitis, food

CC hypersensitivity, urticaria, an immunologic disease of the lung,

CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity

CC pneumonitis, a transplantation associated disease, graft rejection or

CC graft-versus-host disease. The present sequence represents a PRO protein

CC of the invention.

XX Sequence 311 AA;

SQ Query Match 100.0%; Score 73; DB 8; Length 311;

Best Local Similarity 100.0%; Pred. No. 0.00083;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYRCRGDDSKVQE 13
Db 178 NYRCRGDDSKVQE 190

RESULT 6
ABB99112
ID ABB99112 standard; protein; 312 AA.
XX
AC ABB99112;
XX
DT 29-OCT-2002 (first entry)
XX
DE Rat phosphatase #2.
XX
KW Phosphatase; prenilyphosphoric acid; prenily alcohol; rat.
XX
OS Rattus norvegicus.
XX
PN WO200253751-A1.
XX
PD 11-JUL-2002.
XX
PF 20-DEC-2001; 2001WO-JP011223.
XX
PR 28-DEC-2000; 2000JP-00401515.
PR 28-DEC-2000; 2000JP-00401806.
PA (TOYW) TOYOTA CHUO KENKUTUSHO KK.
PA (TOYT) TOYOTA JIDOSHA KK.
XX
PI Tokuhiko K, Muramoto N, Yamada Y, Asami O, Hirai M, Ohto C;
PI Obata S, Muramatsu M;
XX
XX WPI; 2002-643303/69.
DR N-PSDB; ABQ78948.
XX
PT Phosphatase polypeptides and encoded polynucleotides with substrate-
PT specificity for prenilyphosphoric acid, applicable in mass-production of
PT all-trans-prenily alcohols by culturing transformants.
XX
PS Claim 1; Page 67-68; 93pp; Japanese.
XX
CC The invention relates to a novel phosphatase polypeptide, exhibiting
CC substrate-specificity for not less than 15C prenilyphosphoric acid. The
CC polynucleotides and encoded polypeptides are applicable in mass-
CC production of all-trans-prenily alcohols. The sequence represents a rat
CC phosphatase polypeptide of the invention
XX
SQ Sequence 312 AA;

Query Match 94.5%; Score 69; DB 5; Length 312;
Best Local Similarity 92.3%; Pred. No. 0.0039;
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYRCRGDDSKVQE 13
Db 179 NYRCRGDDSKVQE 191

RESULT 7
ADE60701
ID ADE60701 standard; protein; 312 AA.
XX
AC ADE60701;
XX
DT 29-JAN-2004 (first entry)
XX
DE Rat Protein CAA69106, SEQ ID NO 6613.
XX

KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW Chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS Rattus norvegicus.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woelf C, D'urso D, Befort K, Costigan M;
XX
XX WPI; 2003-268312/26.
DR GENBANK; CAA69106.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 312 AA;

Query Match 94.5%; Score 69; DB 7; Length 312;
Best Local Similarity 92.3%; Pred. No. 0.0039;
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYRCRGDDSKVQE 13
Db 179 NYRCRGDDSKVQE 191

RESULT 8
ADB74255
ID ADB74255 standard; protein; 705 AA.
XX
XX ADB74255;
XX

XX	Key	Location/Qualifiers
FH	Peptide	1. .22
FT		/label= Signal_peptide
FT	Protein	23. 1304
FT		/label= Mature_protein
XX	WO2003004524-A2.	
PN	16-JAN-2003.	
XX		
PD	02-JUL-2002; 2002WO-GB003045.	
XX		
PF	02-JUL-2001; 2001GB-00016185.	
XX		
PR	(IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.	
PA		
XX	Claudianos C, Crompton TK, Dessens JT, Sinden RE, Trueman HE;	
PI	WPI; 2003-210339/20.	
XX	N-PSDB; ABZ58196.	
DR		
XX	New secreted proteins from malarial parasites (e.g. Plasmodium falciparum), useful for detecting or diagnosing malaria, or as a vaccine or medicament for the treatment and/or prophylaxis of malaria.	
PT		
PT	Claim 1; Fig 1; 67pp; English.	
XX		
PS	The present sequence is the protein sequence of novel secreted protein PbSR from the mouse malaria parasite, Plasmodium berghei. The invention provides novel modular secreted proteins from Plasmodium falciparum (PfSR), P. berghei (PbSR) and Plasmodium yoelii (PySR) that are post-translationally modified in sporozoites. These include Limulus clotting factor, lipid binding, scavenger receptor cysteine-rich and pentraxin domains. Homologues were identified in Plasmodium knowlesi (PkSR), Plasmodium chabaudi (PcSR), Cryptosporidium parvum (CpSR), Toxoplasma gondii (TgSR) and Theileria parva (TpSR). PbSR, PfSR and PySR are from a unique and previously unrecognized family containing domains associated with binding and modulating host proteins involved in immunity. They are the first proteins from Plasmodium which are secreted and which show a gene disruption phenotype in both vertebrate and mosquito, making them highly useful for the production of malaria vaccines. Transgenic PbSR null mutants have attenuated growth in immunocompetent mice. They fail to produce sporozoites in Anopheles stephensi mosquitoes and consequently are dead-end parasites in the midgut of the mosquito. The novel Plasmodium proteins, and the nucleic acids encoding them, are used in claimed methods for the detection/diagnosis of malaria, and are also used in vaccine compositions. Agents capable of antagonising, inhibiting or interfering with the function or expression of the secreted protein are used in the treatment and/or prophylaxis of malaria	
XX	Sequence 1304 AA;	
XX		
XX	Query Match	57.5%; Score 42; DB 6; Length 1304;
XX	Best Local Similarity	41.7%; Pred. No. 5.2e+02;
XX	Matches	5; Conservative 6; Mismatches 1; Indels 0; Gaps
QY	1 NYRCRGDSDSKVQ 12	
	:: :: :: ::	
DB	614 NFRCKRGDEANLK 625	
RESULT 10		
ABG23434		
ID	ABG23434 standard; protein; 152 AA.	
XX	ABG23434;	
XX		
DT	18-FEB-2002 (first entry)	
XX		
DE	Novel human diagnostic protein #23425.	
XX	Human; chromosome mapping; gene mapping; gene therapy; forensic;	

KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

XX WO200175067-A2.

PD 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR N-PSDB; AAS87621.

XX New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.

XX Claim 20; SEQ ID NO 53793; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic amino acid sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 152 AA;

Query Match 56.2%; Score 41; DB 4; Length 152;

Best Local Similarity 70.0%; Pred. No. 85;

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NTRCRGDDSK 10

Db 68 NTRCRGDLTK 77

RESULT 11

ID AAU28253

XX AAU28253 standard; protein; 237 AA.

AC AAU28253;

XX 18-DEC-2001 (first entry)

XX Novel human secretory protein, Seq ID No 610.

KW Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;

KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;

KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;

KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;

KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;

KW gut protection; lung; liver fibrosis; immune deficiency; infection;

KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;

KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;

KW fertility; analgesic; pain; antigen.

XX Homo sapiens.

XX WO200166689-A2.

XX 13-SEP-2001.

XX 05-MAR-2001; 2001WO-US004942.

XX 07-MAR-2000; 2000US-00519705.

PR 19-MAY-2000; 2000US-00574454.

PR 17-JUN-2000; 2000US-00596193.

PR 14-JUL-2000; 2000US-00616847.

PR 19-SEP-2000; 2000US-00665363.

PR 20-OCT-2000; 2000US-00693267.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;

PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;

XX WPI; 2001-589934/66.

DR N-PSDB; AAS45153.

XX Novel polypeptides and nucleic acids obtained from cDNA libraries prepared from various human tissues, for diagnosis and treatment of cancer, neurological, inflammatory, and autoimmune disorders.

XX Example 2; SEQ ID NO 610; 107pp; English.

CC The invention relates to novel isolated human secreted polypeptides (I) and polynucleotides (II). (I) and (II) are useful for treating inflammatory conditions such as arthritis, nephritis, Crohn's disease, ischaemia-reperfusion injury, shock, sepsis, immune responses, and is involved in increasing haematopoiesis, stem cell survival, bone growth and remodeling. (I), (II) and modulators of (II) are useful for prophylaxis or treatment of one or more cancers. (II) is also useful for creating transgenic animals useful for studying the in vivo activities of the polypeptide as well as for studying modulators of the polypeptides. (I) induces the proliferation of neural cells and regeneration of nerve and brain tissue and is useful for the treatment of central and peripheral nervous system diseases and neuropathies, such as Alzheimer's, Parkinson's disease, Huntington's disease, and amyotrophic lateral sclerosis. In addition, (I) is involved in chemotactic or chemokinetic activity, regulation of haematopoiesis and is useful for treating myeloid or lymphoid cell disorders, platelet disorders such as thrombocytopenia and for regeneration of bone, cartilage, tendon, ligament and/or nerve tissue growth, and in tissue repair, healing of burns, incisions, ulcers, for treating osteoporosis, osteoarthritis, bone degenerative disorders, or periodontal disease. Furthermore, (I) is also useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, various immune deficiencies and disorders including severe combined immunodeficiency (SCID), bacterial or fungal infections, autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic reactions and conditions, such as asthma or other respiratory problems. In addition, (I) affects biorhythms or circadian cycles of rhythms, fertility, metabolism, catabolism, anabolism, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides analgesic effects or other pain reducing effects, immunoglobulin like activity and can act as an antigen in a vaccine composition to raise an immune response. AAU28202-AAU28395 represent novel human secreted protein amino acid sequences of the invention

XX Sequence 237 AA;

Query Match 56.2%; Score 41; DB 4; Length 237;

Best Local Similarity 70.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
Db 5 NHRCRGDLTK 14

RESULT 12

ID ADS12199 standard; protein; 237 AA.

XX AC ADS12199;

XX DT 16-DEC-2004 (first entry)

XX DE Human therapeutic contig protein - SEQ ID 2436.

XX KW antiinflammatory; neuroprotective; antianaemic; cytostatic; vulnerary;
XX KW inflammatory; haematopoiesis; immunity; neurodegenerative; stem cell;
XX KW aplastic anaemia; cancer; wound healing; gene therapy.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT Misc-difference 1..237 /label= Unknown, OTHER
XX FT /note= "OTHER = In-frame STOP codon"

XX PN WO2004080148-A2.

XX PD 23-SEP-2004.

XX PF 30-SEP-2003; 2003WO-US030720.

XX PR 02-OCT-2002; 2002US-0416186P.

XX PA (NUVE-) NUVELO INC.

XX PI Tang YT, Auandi V, Ren F, Zhang J, Wehrman T, Wang Z, Ma Y,
XX PI Wang D, Chen R, Zhao QA, Wang J, Ghosh M, Xue AJ, Weng G, Zhou P;
XX DR WPI; 2004-668857/65.
XX DR N-PSDB; ADS11601.

XX PT New polynucleotide, useful in preparing a composition for diagnosing or
XX PT treating inflammatory, neurodegenerative or stem cell disorders, e.g.,
XX PT aplastic anemia or cancer for promoting wound healing.

XX PS Example 2; SEQ ID NO 2436; 718pp; English.

XX CC The invention relates to a novel isolated polynucleotide and the encoded
XX CC polypeptide. The molecules of the invention demonstrate antiinflammatory,
XX CC neuroprotective, antianaemic, cytostatic and vulnerary activities and may
XX CC be useful in preparing a composition for diagnosing or treating
XX CC inflammatory, haematopoietic, immune, neurodegenerative or stem cell
XX CC disorders, such as aplastic anaemia or cancer, as well as for promoting
XX CC wound healing. The molecules may also be utilised during gene therapy
XX CC procedures. The current sequence is that of a human therapeutic contig
XX CC protein of the invention.

XX SQ Sequence 237 AA;

Query Match 56.2%; Score 41; DB 8; Length 237;
Best Local Similarity 70.0%; Pred. No. 1.3e+02;

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
Db 5 NHRCRGDLTK 14

RESULT 13

Claim 2; SEQ ID NO 12; 342pp; English.

ADJ50947
ID ADJ50947 standard; protein; 421 AA.

XX AC ADJ50947;

XX DT 06-MAY-2004 (first entry)

XX DE Human novel protein NOV4b.

XX KW Human; NOVX; autoimmune disease; Alzheimer's disease; stroke; allergy;
XX KW Parkinson's disease; Huntington's disease; multiple sclerosis; anxiety;
XX KW pain; diabetes; graft versus host disease; pancreatitis; obesity; ulcer;
XX KW anaemia; cancer; viral infection; bacterial infection;
XX KW parasitic infection.

XX OS Homo sapiens.

XX PN US2004030096-A1.

XX PD 12-FEB-2004.

XX PF 01-AUG-2002; 2002US-00210281.

XX PR 02-AUG-2001; 2001US-0309501P.

XX PR 03-AUG-2001; 2001US-0310291P.

XX PR 08-AUG-2001; 2001US-0310951P.

XX PR 09-AUG-2001; 2001US-0311292P.

XX PR 13-AUG-2001; 2001US-0311979P.

XX PR 14-AUG-2001; 2001US-0312203P.

XX PR 17-AUG-2001; 2001US-0313201P.

XX PR 20-AUG-2001; 2001US-0313643P.

XX PR 21-AUG-2001; 2001US-0313702P.

XX PR 21-AUG-2001; 2001US-0314031P.

XX PR 23-AUG-2001; 2001US-0314466P.

XX PR 28-AUG-2001; 2001US-0315403P.

XX PR 29-AUG-2001; 2001US-0315853P.

XX PR 05-MAR-2002; 2002US-0361775P.

XX PR 05-MAR-2002; 2002US-0361832P.

XX (GORM/) GORMAN L.

XX (ZERH/) ZERHUSEN B D.

XX (EDIN/) EDINGER S R.

XX (PADI/) PADIGARU M.

XX (GUOX/) GUO X.

XX (KEKU/) KEKUDA R.

XX (ZHON/) ZHONG M.

XX (PATT/) PATTURAJAN M.

XX (MILL/) MILLER C E.

XX (JIWW/) JI W.

XX (PENA/) PENA C E A.

XX (BURG/) BURGESS C E.

XX (SCIO/) SCIORE P.

XX (STON/) STONE D J.

XX (TAUP/) TAUPIER R J.

XX (CASM/) CASMAN S J.

XX (ROTH/) ROTHENBERG M E.

XX (MALY/) MALYANKAR U M.

XX (BOLD/) BOLDG F L.

XX Gorman L, Zerhusen BD, Edinger SR, Padigaru M, Guo X, Kekuda R;
XX Zhong M, Patturajan M, Miller CE, Ji W, Pena CEA, Burgess CE;
XX Sciore P, Stone DJ, Taupier RJ, Casman SJ, Rothenberg ME;
XX Malyankar UM, Boldog FL;
XX WPI; 2004-168942/16.
XX DR N-PSDB; ADJ50946.

XX New NOVX polypeptides and polynucleotides, useful in diagnosing, treating
XX or preventing diseases or conditions, e.g. autoimmune disease,
XX Alzheimer's disease, diabetes, graft versus host disease, cancer or viral
XX or bacterial infections.

XX The invention relates to an isolated NOVX polypeptide (of 44 disclosed)
CC comprising its mature form, a sequence having at least 95% sequence
CC identity to NOVX or a sequence comprising one or more conservative
CC substitutions in the amino acid sequence of NOVX. Also included are a
CC composition comprising NOVX and a carrier, a kit comprising, in one or
CC more containers, the composition, a method of identifying an agent that
CC binds to NOVX, a method for identifying a potential therapeutic agent for
CC use in treatment of a pathology related to aberrant expression or
CC aberrant physiological interactions of NOVX, a method for screening for a
CC modulator of activity of or of latency or predisposition to a pathology
CC associated with NOVX, a method for modulating the activity of NOVX, a
CC method of treating or preventing a pathology associated with NOVX or a
CC pathological state in a mammal, an isolated nucleic acid molecule
CC encoding a NOVX protein, a vector comprising the nucleic acid molecule,
CC a cell comprising the vector, an antibody that immunospecifically binds
CC to NOVX, a method for determining the presence or amount of NOVX or the
CC nucleic acid molecule in a sample, a method for determining the presence
CC of or predisposition to a disease associated with altered levels of
CC expression of NOVX or the nucleic acid molecule in a first mammalian
CC subject and a method of producing NOVX (comprising culturing the cell
CC under conditions that lead to expression of the polypeptide). NOVX is
CC useful in the manufacture of a medicament for treating a syndrome
CC associated with a human disease associated with NOVX. The polypeptides
CC and nucleic acid molecules are useful in diagnosing, treating or
CC preventing diseases or conditions, e.g. autoimmune disease, Alzheimer's
CC disease, stroke, allergies, Parkinson's disease, Huntington's diseases,
CC multiple sclerosis, anxiety, pain, diabetes, graft versus host disease,
CC pancreatitis, obesity, ulcers, anaemia, cancer, viral or bacterial and
CC parasitic infections (many more diseases and disorders are listed in the
CC specification). The present sequence represents a NOVX protein.

XX SQ Sequence 421 AA;

Query Match 56.2%; Score 41; DB 8; Length 421;
Best Local Similarity 70.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 NVYRCGDDSK 10
|:|||||:
Db 197 NHRRCGDLTK 206

RESULT 14
ADJ50945
ID ADJ50945 standard; protein; 424 AA.

XX AC ADJ50945;

XX DT 06-MAY-2004 (first entry)

XX DE Human novel protein NOV4a.

XX Human; NOVX; autoimmune disease; Alzheimer's disease; stroke; allergy;
KW Parkinson's disease; Huntington's disease; multiple sclerosis; anxiety;
KW pain; diabetes; graft versus host disease; pancreatitis; obesity; ulcer;
KW anaemia; cancer; viral infection; bacterial infection;
KW parasitic infection.

XX OS Homo sapiens.

XX FN US2004030096-A1.

XX PD 12-FEB-2004.

XX PF 01-AUG-2002; 2002US-00210281.

XX PR 02-AUG-2001; 2001US-0309501P.

PR 03-AUG-2001; 2001US-0310291P.

PR 08-AUG-2001; 2001US-0310951P.

PR 09-AUG-2001; 2001US-0311292P.

PR 13-AUG-2001; 2001US-0311799P.

PR 14-AUG-2001; 2001US-0312203P.

PR 17-AUG-2001; 2001US-0313201P.
PR 20-AUG-2001; 2001US-0313643P.
PR 20-AUG-2001; 2001US-0313702P.
PR 21-AUG-2001; 2001US-0314031P.
PR 23-AUG-2001; 2001US-0314466P.
PR 28-AUG-2001; 2001US-0315403P.
PR 29-AUG-2001; 2001US-0315853P.
PR 05-MAR-2002; 2002US-0361775P.
PR 05-MAR-2002; 2002US-0361832P.

XX (GORM/) GORMAN L.
PA (ZERR/) ZERRHUSEN B D.
PA (EDIN/) EDINGER S R.
PA (PADI/) PADIGARU M.
PA (GUOX/) GUO X.
PA (KEKU/) KEKUDA R.
PA (ZHON/) ZHONG M.
PA (PATT/) PATTURAJAN M.
PA (MILL/) MILLER C E.
PA (JIWW/) JI W.
PA (PENA/) PENA C E A.
PA (BURG/) BURGESS C E.
PA (SCIO/) SCIORE P.
PA (STON/) STONE D J.
PA (TAUP/) TAUPIER R J.
PA (CASM/) CASMAN S J.
PA (ROTH/) ROTHENBERG M E.
PA (MALY/) MALYANKAR U M.
PA (BOLD/) BOLDG F L.

XX Gorman L, Zerrhuseen BD, Edinger SR, Padigaru M, Guo X, Kekuda R;
PI Zhong M, Patturajan M, Miller CE, Ji W, Pena CE, Burgess CB;
PI Sciore P, Stone DJ, Taupier RJ, Casman SJ, Rothenberg ME;
PI Malyankar UM, Boldog FL;

XX WPI: 2004-168942/16.
DR N-PSDB; ADJ50944.

XX New NOVX polypeptides and polynucleotides, useful in diagnosing, treating
PT or preventing diseases or conditions, e.g. autoimmune disease,
PT Alzheimer's disease, diabetes, graft versus host disease, cancer or viral
PT or bacterial infections.

PS Claim 2; SEQ ID NO 10; 342pp; English.

XX The invention relates to an isolated NOVX polypeptide (of 44 disclosed)
CC comprising its mature form, a sequence having at least 95% sequence
CC identity to NOVX or a sequence comprising one or more conservative
CC substitutions in the amino acid sequence of NOVX. Also included are a
CC composition comprising NOVX and a carrier, a kit comprising, in one or
CC more containers, the composition, a method of identifying an agent that
CC binds to NOVX, a method for identifying a potential therapeutic agent for
CC use in treatment of a pathology related to aberrant expression or
CC aberrant physiological interactions of NOVX, a method for screening for a
CC modulator of activity of or of latency or predisposition to a pathology
CC associated with NOVX, a method for modulating the activity of NOVX, a
CC method of treating or preventing a pathology associated with NOVX or a
CC pathological state in a mammal, an isolated nucleic acid molecule
CC encoding a NOVX protein, a vector comprising the nucleic acid molecule,
CC a cell comprising the vector, an antibody that immunospecifically binds
CC to NOVX, a method for determining the presence or amount of NOVX or the
CC nucleic acid molecule in a sample, a method for determining the presence
CC of or predisposition to a disease associated with altered levels of
CC expression of NOVX or the nucleic acid molecule in a first mammalian
CC subject and a method of producing NOVX (comprising culturing the cell
CC under conditions that lead to expression of the polypeptide). NOVX is
CC useful in the manufacture of a medicament for treating a syndrome
CC associated with a human disease associated with NOVX. The polypeptides
CC and nucleic acid molecules are useful in diagnosing, treating or
CC preventing diseases or conditions, e.g. autoimmune disease, Alzheimer's
CC disease, stroke, allergies, Parkinson's disease, Huntington's diseases,
CC multiple sclerosis, anxiety, pain, diabetes, graft versus host disease,
CC pancreatitis, obesity, ulcers, anaemia, cancer, viral or bacterial and

CC parasitic infections (many more diseases and disorders are listed in the
CC specification). The present sequence represents a NOVX protein.

XX
SQ Sequence 424 AA;

Query Match 56.2%; Score 41; DB 8; Length 424;
Best Local Similarity 70.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
|:|||||:|
Db 198 NHRCRGDLTK 207

RESULT 15
ADA83857
ID ADA83857 standard; protein; 428 AA.

XX ADA83857;

XX 20-NOV-2003 (first entry)

XX Human POM18 protein.

XX human; marker; expressed sequence tag; EST; arabidopsis; tumour;
KW stress-induced phenotype; hyperosmotic stress; colon cancer; immunogen;
KW vaccine.

XX Homo sapiens.

XX WO2002103028-A2.

XX 27-DEC-2002.

XX 30-MAY-2002; 2002WO-IB004189.

XX 30-MAY-2001; 2001US-0293999P.

PR 22-OCT-2001; 2001US-0330457P.

PR 19-FEB-2002; 2002US-0357144P.

XX (BIOM-) BIOMEDICAL CENT.

XX Baranova AV, Yankovsky NK, Kozlov AP, Lobashev AV, Krukovskaya LL;

DR WPI; 2003-175241/17.

DR N-PSDB; ADA83856.

XX Determining if a nucleic acid is a marker for a phenotype/cell type of
PT interest, by global comparison of expressed sequence tags known to be
PT expressed in the phenotype/cell type with all ESTs expressed in normal
PT tissue.

XX Claim 29; Page 269-271; 516pp; English.

XX The invention relates to a novel method for determining if a nucleic acid
CC is a marker for a predetermined phenotype/cell type of interest from a
CC biological species. The method comprises performing a global comparison
CC of a group of expressed sequence tags (ESTs) known to be expressed in the
CC phenotype/cell type of interest with all ESTs expressed in normal tissue
CC in order to identify ESTs that are preferentially expressed in the
CC phenotype/cell of interest. A method of the invention is useful for
CC determining whether a nucleic acid is a marker for a predetermined
CC phenotype or cell type of interest from a biological species, preferably
CC Arabidopsis or human. The cell type of interest is an abnormal cell such
CC as a tumour cell, and the predetermined phenotype is a stress-induced
CC phenotype such as hyperosmotic stress or high salt conditions. A method
CC of the invention is also useful for determining the progression of colon
CC cancer in a human, for detecting a tumour cell, and for regulating or
CC preventing the growth of a tumour cell. An antibody of the invention is
CC useful for detecting the absence or presence of peptides encoded by
CC tumour-associated markers. A polypeptide of the invention is useful as an
CC immunogen for vaccinating an animal. The present sequence represents a
CC tumour-associated antigen of the invention.

XX
SQ Sequence 428 AA;

Query Match 56.2%; Score 41; DB 6; Length 428;
Best Local Similarity 70.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
|:|||||:|
Db 197 NHRCRGDLTK 206

RESULT 16
ABB67864
ID ABB67864 standard; protein; 451 AA.

XX ABB67864;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 30384.

XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-658860/75.

DR N-PSDB; ABL11967.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.

XX Disclosure; SEQ ID NO 30384; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 451 AA;

Query Match 56.2%; Score 41; DB 4; Length 451;
Best Local Similarity 63.6%; Pred. No. 2.6e+02;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 YRCRGDDSKVQ 12
|||||:|
Db 3 YRCRGNESSDQ 13

RESULT 17
ADR09234

ID ADR09234 standard; protein; 481 AA.
AC ADR09234;
XX
DT 04-NOV-2004 (first entry)
XX
DE Human protein useful for treating neurological disease Seq 2740.
XX
KW human; oligo-capping method; diagnostic marker; gene therapy;
KW osteoporosis; neurological disease; Alzheimer's disease;
KW Parkinson's disease; dementia; short memory; cancer;
KW sense or motor function; emotional reaction; fear response; panic;
KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;
KW tranquiliser.
XX
OS Homo sapiens.
XX
PN EP1447413-A2.
XX
PD 18-AUG-2004.
XX
PF 12-FEB-2004; 2004EP-00003145.
XX
PR 14-FEB-2003; 2003JP-00102207.
PR 09-MAY-2003; 2003JP-00131452.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;
PI Wakamatsu A, Ishii S, Nagai K, Irie R;
XX
DR WPI; 2004-593285/57.
XX
DR N-PSDB; ADR07278.
XX
XX New 1995 cDNA, useful for treating osteoporosis, neurological diseases,
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX
PS Claim 1; SEQ ID NO 2740; 2686pp; English.
XX
CC This invention relates to novel, isolated full length human cDNA
CC molecules and the encoded proteins thereof. Specifically, it refers to
CC cDNA clones obtained by an oligo-capping method, where none of these
CC clones are identical to any known human mRNAs. The present invention
CC describes an immunoassay to identify agonists and antagonists, as well as
CC antibodies, antisense molecules and siRNAs that can all be used to bind
CC to and modulate expression of the cDNA molecules. As such, these
CC molecules are useful for diagnostic markers or therapeutic targets for
CC the various diseases or morbid states. In particular, they are useful in
CC gene therapy for treating osteoporosis, neurological disease, Alzheimer's
CC disease, Parkinson's disease, dementia, short memory and various cancers,
CC as well as for maintaining equilibrium of sense or motor function, and
CC for treating emotional reaction, fear response and panic. Accordingly,
CC they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,
CC cytostatic and tranquiliser activities. This polypeptide is a protein
CC encoded by a full length human cDNA sequence of the invention. NOTE: This
CC sequence is not given in the sequence listing of the specification but
CC can be obtained on CD-ROM from the European Patent Office, Vienna Sub-
CC office.
XX
SQ Sequence 481 AA;

Query Match 56.2%; Score 41; DB 8; Length 481;
Best Local Similarity 70.0%; Pred. No. 2.8e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDSDK 10
|:|||||:
Db 250 NHRCRGDLTK 259

RESULT 18
ADSI0966
ID ADSI0966 standard; protein; 481 AA.

XX ADSI0966;
AC
XX
DT 16-DEC-2004 (first entry)
XX
DE Human therapeutic protein - SEQ ID 1203.
XX
KW antiinflammatory; neuroprotective; antianaemic; cytostatic; vulnerary;
KW inflammatory; haematopoiesis; immunity; neurodegenerative; stem cell;
KW aplastic anaemia; cancer; wound healing; gene therapy.
XX
OS Homo sapiens.
XX
PN WO2004080148-A2.
XX
XX
PD 23-SEP-2004.
XX
PP 30-SEP-2003; 2003WO-US030720.
XX
PR 02-OCT-2002; 2002US-0416186P.
XX
PA (NUVE-) NUVELO INC.
XX
PI Tang YT, Asundi V, Ren F, Zhang J, Wehrman T, Wang Z, Ma Y;
PI Wang D, Chen R, Zhao QA, Wang J, Ghosh M, Xue AJ, Weng G, Zhou P;
XX
XX WPI; 2004-668857/65.
DR N-PSDB; ADSI0282.
XX
XX New polynucleotide, useful in preparing a composition for diagnosing or
PT treating inflammatory, neurodegenerative or stem cell disorders, e.g.,
PT aplastic anemia or cancer for promoting wound healing.
XX
XX Claim 20; SEQ ID NO 1203; 718pp; English.
XX
CC The invention relates to a novel isolated polynucleotide and the encoded
CC polypeptide. The molecules of the invention demonstrate antiinflammatory,
CC neuroprotective, antianaemic, cytostatic and vulnerary activities and may
CC be useful in preparing a composition for diagnosing or treating
CC inflammatory, haematopoietic, immune, neurodegenerative or stem cell
CC disorders, such as aplastic anaemia or cancer, as well as for promoting
CC wound healing. The molecules may also be utilised during gene therapy
CC procedures. The current sequence is that of a human therapeutic protein
CC of the invention. The current sequence is not shown explicitly within the
CC specification but can be accessed from the WIPO web-site.
XX
SQ Sequence 481 AA;

Query Match 56.2%; Score 41; DB 8; Length 481;
Best Local Similarity 70.0%; Pred. No. 2.8e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDSDK 10
|:|||||:
Db 250 NHRCRGDLTK 259

RESULT 19
ADM04303
ID ADM04303 standard; protein; 489 AA.
XX
AC ADM04303;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human protein of the invention SEQ ID NO:2988.
XX
KW human; gene therapy; diagnostic marker; pharmaceutical.
XX
OS Homo sapiens.
XX
PN EP1347046-A1.
XX

CC wide range of psychiatric disorders; and (xi) to treat brain or lung
CC injury associated with protease expression (specifically emphysema or
CC bronchitis)
XX
SQ Sequence 875 AA;

Query Match 56.2%; Score 41; DB 2; Length 875;
Best Local Similarity 54.5%; Pred. No. 5.1e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 NVRCRGDDSKV 11
| | | | | :
Db 236 NVRCRGDEENI 246

RESULT 22
ADQ21226
ID ADQ21226 standard; protein; 875 AA.
XX
AC ADQ21226;
XX
DT 26-AUG-2004 (first entry)
XX
DE Human soft tissue sarcoma-upregulated protein - SEQ ID 4046.
XX
KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.
XX
OS Homo sapiens.
XX
PN WO2004048938-A2.
XX
PD 10-JUN-2004.
XX
PF 26-NOV-2003; 2003WO-US038193.
XX
PR 26-NOV-2002; 2002US-0429739P.
XX
PA (PROT-) PROTEIN DESIGN LABS INC.
XX
PI Aziz N, Ginsburg WM, Zlotnik A;
XX
WPI; 2004-441208/41.
XX
PT Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX
PS Example 2; SEQ ID NO 4046; 210pp; English.
XX
CC The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cytostatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC protein of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.
XX
SQ Sequence 875 AA;

Query Match 56.2%; Score 41; DB 8; Length 875;
Best Local Similarity 54.5%; Pred. No. 5.1e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 NVRCRGDDSKV 11
| | | | | :
Db 236 NVRCRGDEENI 246

RESULT 23
ADQ89082
ID ADQ89082 standard; protein; 875 AA.
XX
AC ADQ89082;
XX
DT 21-OCT-2004 (first entry)
XX
DE Human urological disorder related protein 14063 SEQ:34.
XX
KW urological disorder; uropathic; cytostatic; urinary incontinence;
KW benign prostatic hyperplasia; human.
XX
OS Homo sapiens.
XX
PN WO2004065576-A2.
XX
PD 05-AUG-2004.
XX
PF 14-JAN-2004; 2004WO-US000750.
XX
PR 15-JAN-2003; 2003US-0440318P.
PR 04-FEB-2003; 2003US-0444783P.
PR 27-MAR-2003; 2003US-0457901P.
PR 08-MAY-2003; 2003US-0468775P.
PR 19-MAY-2003; 2003US-0471614P.
PR 16-JUN-2003; 2003US-0478742P.
PR 18-JUL-2003; 2003US-048529P.
PR 30-JUL-2003; 2003US-0491156P.
PR 02-SEP-2003; 2003US-0499594P.
PR 26-SEP-2003; 2003US-0506332P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Karicheti V, Silos-Santiago I, Eliasof SD;
XX
WPI; 2004-562167/54.
XX
DR N-PSDB; ADQ89081.
XX
PT Use of polypeptides related to urological disorders, e.g. 44390, 54181,
PT 211 or for identifying a compound capable of treating a urological
PT disorder or identifying and treating a subject having a urological
PT disorder.
XX
PS Claim 1; SEQ ID NO 34; 542pp; English.
XX
CC The present invention describes the use of polypeptides related to
CC urological disorders for identifying a compound capable of treating a
CC urological disorder, identifying a subject having a urological disorder,
CC or treating a subject having a urological disorder. Also described: (1) a
CC method for identifying a compound capable of treating a urological
CC disorder; (2) a method for identifying a subject having a urological
CC disorder; and (3) a method for treating a subject having a urological
CC disorder. The compound has uropathic and cytostatic activities. The
CC polypeptides related to urological disorders are useful for identifying a
CC compound capable of treating a urological disorder, identifying a subject
CC having a urological disorder, or treating a subject having a urological
CC disorder. Disorders include urinary incontinence and benign prostatic
CC hyperplasia. The present sequence represents a human urological disorder
CC related protein, which is used in the exemplification of the present
CC invention.
XX
SQ Sequence 875 AA;

Query Match 56.2%; Score 41; DB 8; Length 875;
Best Local Similarity 54.5%; Pred. No. 5.1e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 NVRCRGDDSKV 11
| | | | | :
Db 236 NVRCRGDEENI 246

RESULT 24
 ABP56183
 ID ABP56183 standard; peptide; 20 AA.
 XX
 AC
 XX
 DT 28-MAR-2003 (first entry)
 XX
 DE
 XX
 DE Chimeric peptide sequence #5.
 XX
 KW Mitochondrial membrane permeabilisation; mitochondrion; PTPC; chimeric;
 KW permeability transition pore complex; virucide; neuroprotective;
 KW vasotropic; cytosolic; infection; cell death regulation; apoptosis;
 KW mitochondrial permeability transition pore complex modulator; cancer;
 KW apoptogenic; ischaemia; neurodegenerative disease; fulminant hepatitis.
 XX
 OS Synthetic.
 XX
 PN WO200261105-A2.
 XX
 PD 08-AUG-2002.
 XX
 PF 01-FEB-2002; 2002WO-EP001633.
 XX
 PR 02-FEB-2001; 2001US-0265594P.
 XX
 PA (INSP) INST PASTEUR.
 PA (CNRS) CENT NAT RECH SCT.
 XX
 PI Edelman L, Jacotot E, Briand J;
 XX
 DR WPI; 2002-619260/66.
 XX
 XX New chimeric bifunctional molecules that target specific cells and
 PT regulate the apoptosis function of the permeability transition pore
 PT complex of the mitochondria, useful for treating or preventing e.g.
 PT cancer or ischemia.
 XX
 PS Disclosure; Page 23; 76pp; English.
 XX
 CC The present invention describes a chimeric bifunctional molecule (I)
 CC comprising at least a first functional molecule covalently linked to a
 CC second functional molecule, which is able to modulate the activity of the
 CC permeability transition pore complex (PTPC) of the mitochondria. (I) has
 CC the function of specifically targeting and entering a tissue cell
 CC population. The second functional molecule has the function of
 CC specifically targeting, and inducing or preventing the death of the cells
 CC by apoptosis by regulating the opening or the closing of the PTPC of the
 CC mitochondria or its fragment. (I) has virucide, neuroprotective,
 CC vasotropic and cytostatic activities, and can be used as a mitochondrial
 CC permeability transition pore complex (PTPC) modulator. (I) is useful for
 CC treating or preventing a pathological infection or disease. (I) is also
 CC useful for regulating cell death regulatory molecules, specifically the
 CC apoptogenic function of the PTPC, for treating e.g. cancer, ischaemia,
 CC neurodegenerative diseases, fulminant hepatitis or viral infections. The
 CC present sequence represents a chimeric peptide which is given in the
 CC exemplification of the present invention
 XX
 SQ Sequence 20 AA;
 Query Match 54.8%; Score 40; DB 5; Length 20;
 Best Local Similarity 58.3%; Pred No. 16;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 NYRCRGDSDSKVQ 12
 Db ||||| :
 2 NGRCGGDKRTQ 13
 RESULT 25
 ABB99111
 ID ABB99111 standard; protein; 282 AA.
 XX
 AC ABB99111;
 XX
 DT 29-OCT-2002 (first entry)
 XX
 DE Rat phosphatase #1.
 XX
 KW Phosphatase; prenlyphosphoric acid; prenyl alcohol; rat.
 XX
 OS Rattus norvegicus.
 XX
 PN WO200253751-A1.
 XX
 PD 11-JUL-2002.
 XX
 PF 20-DEC-2001; 2001WO-JP011223.
 XX
 PR 28-DEC-2000; 2000JP-00401515.
 PR 28-DEC-2000; 2000JP-00401806.
 XX
 PA (TOYM) TOYOTA CHUO KENKYUSHO KK.
 PA (TOYT) TOYOTA JIDOSHA KK.
 XX
 PI Tokuhiko K, Muramoto N, Yamada Y, Asami O, Hirai M, Ohto C;
 PI Obata S, Muramatsu M;
 XX
 DR WPI; 2002-643303/69.
 DR N-PSDB; ABQ78947.
 XX
 PT Phosphatase polypeptides and encoded polynucleotides with substrate-
 PT specificity for prenlyphosphoric acid, applicable in mass-production of
 PT all-trans-prenyl alcohols by culturing transformants.
 XX
 PS Claim 1; Page 64-65; 93pp; Japanese.
 XX
 CC The invention relates to a novel phosphatase polypeptide, exhibiting
 CC substrate-specificity for not less than 15C prenlyphosphoric acid. The
 CC polynucleotides and encoded polypeptides are applicable in mass-
 CC production of all-trans-prenyl alcohols. The sequence represents a rat
 CC phosphatase polypeptide of the invention
 XX
 SQ Sequence 282 AA;
 Query Match 54.8%; Score 40; DB 5; Length 282;
 Best Local Similarity 46.2%; Pred No. 2.4e+02;
 Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
 QY 1 NYRCRGDSDSKVQE 13
 Db ||||| :
 150 NFVCGQNEQKVRE 162
 RESULT 26
 ADE58904
 ID ADE58904 standard; protein; 282 AA.
 XX
 AC ADE58904;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Rat Protein NP_071983, SEQ ID NO 4792.
 XX
 KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 KW Chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
 XX
 OS Rattus norvegicus.
 XX
 PN WO2003016475-A2.
 XX
 PD 27-FEB-2003.
 XX
 PF 14-AUG-2002; 2002WO-US025765.

```

XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
XX (GEHO ) GEN HOSPITAL CORP.
PA (FARB ) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
PI WPI; 2003-268312/26.
XX GENBANK; NP_071983.
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (SNI)), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 282 AA;
XX
XX Query Match 54.8%; Score 40; DB 7; Length 282;
XX Best Local Similarity 46.2%; Pred. No. 2.4e+02;
XX Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 NYRCRGDDSKVOE 13
XX |:|:|:|:|:|
XX Db 150 NFVCQNEQKVRE 162
XX
XX RESULT 27
XX ADE58900
XX ID ADE58900 standard; protein; 282 AA.
XX
XX AC ADE58900;
XX
XX XX 29-JAN-2004 (first entry)
XX
XX Rat Protein NP_071983, SEQ ID NO 4788.
XX
XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
XX Rattus norvegicus.
XX
XX WO2003016475-A2.
XX

```

```

XX 27-FEB-2003.
XX
XX 14-AUG-2002; 2002WO-US025765.
XX
XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
XX (GEHO ) GEN HOSPITAL CORP.
PA (FARB ) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
PI WPI; 2003-268312/26.
XX GENBANK; NP_071983.
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (SNI)), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 282 AA;
XX
XX Query Match 54.8%; Score 40; DB 7; Length 282;
XX Best Local Similarity 46.2%; Pred. No. 2.4e+02;
XX Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 NYRCRGDDSKVOE 13
XX |:|:|:|:|:|
XX Db 150 NFVCQNEQKVRE 162
XX
XX RESULT 28
XX ADD45532
XX ID ADD45532 standard; protein; 282 AA.
XX
XX AC ADD45532;
XX
XX XX 29-JAN-2004 (first entry)
XX
XX Rat Protein NP_071983, SEQ ID NO 11196.
XX
XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
XX

```

XX OS Rattus norvegicus.
 XX PN WO2003016475-A2.
 XX PD 27-FEB-2003.
 XX PF 14-AUG-2002; 2002WO-US025765.
 XX PR 14-AUG-2001; 2001US-0312147P.
 XX PR 01-NOV-2001; 2001US-0346382P.
 XX PR 26-NOV-2001; 2001US-0333347P.
 XX (GEO) GEN HOSPITAL CORP.
 XX PA (FARB) BAYER AG.
 XX PI Woolf C, D'urao D, Befort K, Costigan M;
 XX WPI; 2003-268312/26.
 XX DR GENBANK; NP_071983.
 XX PT New composition comprising two or more isolated polypeptides, useful for
 XX PT preparing a medicament for treating pain in an animal.
 XX PS Claim 1; Page; 1017pp; English.
 XX CC The invention discloses a composition comprising two or more isolated rat
 XX CC or human polynucleotides or a polynucleotide which represents a fragment,
 XX CC derivative or allelic variation of the nucleic acid sequence. Also
 XX CC claimed are a vector comprising the novel polynucleotide, a host cell
 XX CC comprising the vector, a method for identifying a nucleotide sequence
 XX CC which is differentially regulated in an animal subjected to pain and a
 XX CC kit to perform the method, an array, a method for identifying an agent
 XX CC that increases or decreases the expression of the polynucleotide sequence
 XX CC that is differentially expressed in neuronal tissue of a first animal
 XX CC subjected to pain, a method for identifying a compound which regulates
 XX CC the expression of a polynucleotide sequence which is differentially
 XX CC expressed in an animal subjected to pain, a method for identifying a
 XX CC compound that regulates the activity of one or more of the
 XX CC polynucleotides, a method for producing a pharmaceutical composition, a
 XX CC method for identifying a compound or small molecule that regulates the
 XX CC activity in an animal of one or more of the polypeptides given in the
 XX CC specification, a method for identifying a compound useful in treating
 XX CC pain and a pharmaceutical composition comprising the one or more
 XX CC polypeptides or their antibodies. The polynucleotide or the compound that
 XX CC modulates its activity is useful for preparing a medicament for treating
 XX CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 XX CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 XX CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 XX CC the specification) which is differentially expressed during pain. Note:
 XX CC The sequence data for this patent did not form part of the printed
 XX CC specification, but was obtained in electronic form directly from WIPO at
 XX CC ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 282 AA;
 Query Match 54.8%; Score 40; DB 7; Length 282;
 Best Local Similarity 46.2%; Pred. No. 2.4e+02;
 Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
 QY 1 NYRCRGDDSKVQVE 13
 | : | : | : | : |
 Db 150 NFVCGNEGQKVE 162
 RESULT 29
 ADS26705
 ID ADS26705 standard; protein; 301 AA.
 XX AC ADS26705;
 XX DT 02-DEC-2004 (first entry)
 XX

DE Bacterial polypeptide #15738.
 XX Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.
 XX Bacteria.
 OS US2003233675-A1.
 XX PN 18-DEC-2003.
 XX PD 20-FEB-2003; 2003US-00369493.
 XX PF 21-FEB-2002; 2002US-0360039P.
 XX PR (CAOY/) CAO Y.
 XX PA (HINK/) HINKLE G J.
 XX PA (SLAT/) SLATER S C.
 XX PA (CHEN/) CHEN X.
 XX PA (GOLD/) GOLDMAN B S.
 XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX WPI; 2004-061375/06.
 XX DR New recombinant DNA construct comprising a promoter positioned to provide
 XX PT for expression of a polynucleotide encoding a polypeptide from a
 XX PT microbial source, useful for producing plants with improved properties.
 XX PS Claim 1; SEQ ID NO 15738; 122pp; English.
 XX CC The invention relates to a recombinant DNA construct comprising a
 XX CC promoter functional in a plant cell, where the promoter is positioned to
 XX CC provide for expression of a polynucleotide encoding a polypeptide from a
 XX CC microbial source. The invention also relates to a transformed plant
 XX CC comprising the recombinant DNA construct and a method of producing a
 XX CC transformed plant having an improved property. The plant is a crop plant
 XX CC such as maize or soybean. The method of producing a transformed plant
 XX CC having an improved property comprises transforming a plant with the
 XX CC recombinant DNA construct and growing the transformed plant, where the
 XX CC polynucleotide or polypeptide is useful for improving plant properties.
 XX CC The recombinant DNA construct is useful for producing plants with
 XX CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 XX CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 XX CC increased resistance to plant disease, better growth rate by modification
 XX CC of the cell cycle pathway with plant growth regulators, increased rate of
 XX CC homologous recombination, modified seed oil or protein yield and/or
 XX CC content, improved yield by modification of carbohydrate, nitrogen or
 XX CC phosphorus use and/or uptake, by modification of photosynthesis or by
 XX CC providing improved plant growth and development under at least one stress
 XX CC condition, improved lignin production or improved galactomannan
 XX CC production. This sequence represents a bacterial polypeptide used in the
 XX CC scope of the invention. Note: The sequence data for this patent did not
 XX CC form part of the printed specification but was obtained in electronic
 XX CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX SQ Sequence 301 AA;
 Query Match 54.8%; Score 40; DB 8; Length 301;
 Best Local Similarity 50.0%; Pred. No. 2.5e+02;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 QY 4 CRGDDSKVQVE 13
 | : | : | : | : |
 Db 200 CRGDDARIHD 209
 RESULT 30

ADS27091
ID ADS27091 standard; protein; 301 AA.
XX
AC ADS27091;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #16124.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 19-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 16124; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 301 AA;
Query Match 54.8%; Score 40; DB 8; Length 301;
Best Local Similarity 50.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 4 CRGDDSKVOE 13
DB 200 CRGDDARIHD 209
RESULT 31
ADS26338
ID ADS26338 standard; protein; 302 AA.
XX
AC ADS26338;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #15371.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 15371; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX

CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 302 AA;

Query Match 54.8%; Score 40; DB 8; Length 302;

Best Local Similarity 50.0%; Pred. No. 2.5e+02;

Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 4 CRGDDSKVQES 13

|||||:::;

Db 203 CRGDDARIHD 212

RESULT 32

ABU31582

ID ABU31582 standard; protein; 400 AA.

XX AC ABU31582;

XX DT 19-JUN-2003 (first entry)

XX DE Protein encoded by Prokaryotic essential gene #17109.

XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX OS Klebsiella pneumoniae.

XX PN WO200277183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-US009107.

XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2001US-00072851.

XX PR 06-MAR-2002; 2002US-0362699P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX DR WPI; 2003-029926/02.

XX DR N-PSDB; ACA35452.

XX PT New antisense nucleic acids, useful for identifying proteins or screening

XX PT for homologous nucleic acids required for cellular proliferation to

XX PT isolate candidate molecules for rational drug discovery programs.

XX PS Claim 25; SEQ ID NO 59506; 1766pp; English.

XX CC The invention relates to an isolated nucleic acid comprising any one of
the 6213 antisense sequences given in the specification where expression
of the nucleic acid inhibits proliferation of a cell. Also included are:

CC (1) a vector comprising a promoter operably linked to the nucleic acid
encoding a polypeptide whose expression is inhibited by the antisense
nucleic acid; (2) a host cell containing the vector; (3) an isolated
polypeptide or its fragment whose expression is inhibited by the
antisense nucleic acid; (4) an antibody capable of specifically binding
the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
proliferation or the activity of a gene in an operon required for

CC proliferation; (7) identifying a compound that influences the activity of
the gene product or that has an activity against a biological pathway
required for proliferation, or that inhibits cellular proliferation; (8)
identifying a gene required for cellular proliferation or the biological
pathway in which a proliferation-required gene or its gene product lies
or a gene on which the test compound that inhibits proliferation of an
organism acts; (9) manufacturing an antibiotic; (10) profiling a
compound's activity; (11) a culture comprising strains in which the gene
product is overexpressed or underexpressed; (12) determining the extent
to which each of the strains is present in a culture or collection of

CC strains; or (13) identifying the target of a compound that inhibits the
proliferation of an organism. The antisense nucleic acids are useful for
identifying proteins or screening for homologous nucleic acids required
for cellular proliferation to isolate candidate molecules for rational
drug discovery programs, or for screening homologous nucleic acids
required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
K. pneumoniae or *P. aeruginosa*. The present sequence is encoded by one of
the target prokaryotic essential genes. Note: The sequence data for this
patent did not form part of the printed specification, but was obtained
in electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 400 AA;

Query Match 54.8%; Score 40; DB 6; Length 400;

Best Local Similarity 85.7%; Pred. No. 3.4e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDD 8

:|||||

Db 39 FRCRGDD 45

RESULT 33

ABU45262

ID ABU45262 standard; protein; 400 AA.

XX AC ABU45262;

XX DT 19-JUN-2003 (first entry)

XX DE Protein encoded by Prokaryotic essential gene #30789.

XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX OS Salmonella paratyphi.

XX PN WO200277183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-US009107.

XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.

XX PR 06-MAR-2002; 2002US-0362699P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX DR WPI; 2003-029926/02.

XX DR N-PSDB; ACA49132.

XX PT New antisense nucleic acids, useful for identifying proteins or screening
for homologous nucleic acids required for cellular proliferation to
isolate candidate molecules for rational drug discovery programs.

XX PS Claim 25; SEQ ID NO 73186; 1766pp; English.

XX CC The invention relates to an isolated nucleic acid comprising any one of
the 6213 antisense sequences given in the specification where expression
of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
encoding a polypeptide whose expression is inhibited by the antisense
nucleic acid; (2) a host cell containing the vector; (3) an isolated
polypeptide or its fragment whose expression is inhibited by the
antisense nucleic acid; (4) an antibody capable of specifically binding
the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
proliferation or the activity of a gene in an operon required for

CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 400 AA;

Query Match 54.8%; Score 40; DB 6; Length 400;
 Best Local Similarity 85.7%; Pred. No. 3.4e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDD 8
 :|||||
 Db 39 FRCRGDD 45

RESULT 34
 ABU48051
 ID ABU48051 standard; protein; 400 AA.
 XX
 AC ABU48051;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #33578.
 XX
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Salmomella typhi.
 XX WO200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US0009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI: 2003-0299926/02.
 DR N-PSDB; ACA51921.
 XX
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 XX Claim 25; SEQ ID NO 75975; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 400 AA;

Query Match 54.8%; Score 40; DB 6; Length 400;
 Best Local Similarity 85.7%; Pred. No. 3.4e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDD 8
 :|||||
 Db 39 FRCRGDD 45

RESULT 35
 ABU15369
 ID ABU15369 standard; protein; 400 AA.
 XX
 AC ABU15369;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #896.
 XX
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX Escherichia coli.
 XX WO200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US0009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.
 DR N-PSDB; ACA19239.
 XX
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 XX Claim 25; SEQ ID NO 43293; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway;
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 400 AA;
 SQ
 Query Match 54.8%; Score 40; DB 6; Length 400;
 Best Local Similarity 85.7%; Pred. No. 3.4e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 YRCRGDD 8
 Db :|||||
 39 FRCRGDD 45
 RESULT 36
 ID ABO65513
 AC ABO65513 standard; protein; 424 AA.
 XX
 AC ABO65513;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Klebsiella pneumoniae polypeptide seqid 12030.
 XX
 KW Recombinant expression vector; transcription regulatory element;
 KW Klebsiella pneumoniae protein; antibacterial; Vaccine.
 XX
 OS Klebsiella pneumoniae.
 XX
 XX US6610836-B1.
 XX
 PD 26-AUG-2003.
 XX
 PF 27-JAN-2000; 2000US-00489039.

XX 29-JAN-1999; 99US-0117747P.
 PR (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PA Breton GL, Osborne M;
 XX
 PI WPI; 2003-895346/82.
 XX
 DR N-PSDB; ACH99064.
 XX
 XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
 PT preparing a vaccine composition against Klebsiella pneumoniae.
 PT
 PS Disclosure; SEQ ID NO 12030; 932pp; English.
 XX
 CC The invention describes a new isolated nucleic acid encoding a Klebsiella
 CC pneumoniae polypeptide. Also described are: a recombinant expression
 CC vector comprising the nucleic acid, operably linked to a transcription
 CC regulatory element; and a cell comprising the recombinant expression
 CC vector. The nucleic acid is useful for preparing a vaccine composition
 CC against Klebsiella pneumoniae. This is the amino acid sequence of a
 CC Klebsiella pneumoniae polypeptide of the invention
 XX
 SQ Sequence 424 AA;
 Query Match 54.8%; Score 40; DB 7; Length 424;
 Best Local Similarity 85.7%; Pred. No. 3.6e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 YRCRGDD 8
 Db :|||||
 63 FRCRGDD 69
 RESULT 37
 ID ABO65513 standard; protein; 1299 AA.
 XX
 AC ABO65513;
 XX
 DT 22-APR-2003 (first entry)
 XX
 DE Plasmodium yoelii modular secreted protein PySR.
 XX
 KW PfSR; protein secretion; malaria; diagnosis; vaccine; antimalarial;
 KW protozoacide.
 XX
 OS Plasmodium yoelii.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..22
 FT /label= signal_peptide
 FT Protein 23..2399
 FT /label= Mature_protein
 FT Misc-difference 175
 FT /note= "encoded by ATK"
 XX
 PN WO2003004524-A2.
 XX
 PD 16-JAN-2003.
 XX
 PF 02-JUL-2002; 2002WO-GB003045.
 XX
 PR 02-JUL-2001; 2001GB-00016185.
 XX
 PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
 XX
 PI Claudianos C, Crompton TK, Dessens JT, Sindén RE, Trueman HF;
 XX
 DR WPI; 2003-210339/20.
 DR N-PSDB; ABZ58198.
 XX
 PF New secreted proteins from malarial parasites (e.g. Plasmodium

PT falciparum), useful for detecting or diagnosing malaria, or as a vaccine
PT or medicament for the treatment and/or prophylaxis of malaria.

XX
XX Claim 1; Fig 1; 67pp; English.

XX
XX The present sequence is the protein sequence of novel secreted protein
CC PySR from the mouse malaria parasite, Plasmodium yoelii. The invention
CC provides novel modular secreted proteins from Plasmodium falciparum
CC (PfSR), Plasmodium berghei (PbSR) and P. yoelii (PySR) that are post-
CC translationally modified in sporozoites. These include Limulus clotting
CC factor, lipid binding, scavenger receptor cysteine-rich and pentraxin
CC domains. Homologues were identified in Plasmodium knowlesi (PkSR),
CC Plasmodium chabaudi (PcSR), Cryptosporidium parvum (CpSR), Toxoplasma
CC gondii (TgSR) and Theileria parva (TpSR). PbSR, PfSR and PySR are from a
CC unique and previously unrecognized family containing domains associated
CC with binding and modulating host proteins involved in immunity. They are
CC the first proteins from Plasmodium which are secreted and which show a
CC gene disruption phenotype in both vertebrate and mosquito, making them
CC highly useful for the production of malaria vaccines. Transgenic PbSR
CC null mutants have attenuated growth in immunocompetent mice. They fail to
CC produce sporozoites in Anopheles stephensi mosquitoes and consequently
CC are dead-end parasites in the midgut of the mosquito. The novel
CC Plasmodium proteins, and the nucleic acids encoding them, are used in
CC claimed methods for the detection/diagnosis of malaria, and are also used
CC in vaccine compositions. Agents capable of antagonising, inhibiting or
CC interfering with the function or expression of the secreted protein are
CC used in the treatment and/or prophylaxis of malaria

XX SQ Sequence 1299 AA;

Query Match 54.8%; Score 40; DB 6; Length 1299;

Best Local Similarity 41.7%; Pred. No. 1.1e+03;

Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 1 NYRCRGDSKVQ 12

Db 614 NFRCKGDEPNLK 625

RESULT 38

ABB04096

ID ABB04096 standard; protein; 74 AA.

XX AC ABB04096;

XX DT 08-JAN-2002 (first entry)

XX DE Human musculoskeletal system related polypeptide SEQ ID NO 2043.

XX KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein;
KW musculoskeletal system.

XX OS Homo sapiens.

XX PN WO200155367-A1.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001138.

XX PR 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229143P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234224P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.

PR 20-OCT-2000; 2000US-02411786P.
 PR 20-OCT-2000; 2000US-02411787P.
 PR 20-OCT-2000; 2000US-02411808P.
 PR 20-OCT-2000; 2000US-02411809P.
 PR 20-OCT-2000; 2000US-02411826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 PI WPI; 2001-451937/48.
 XX N-PSDB; AAL35678.
 DR
 XX
 XX Isolated polypeptide for treating, preventing and/or prognosing
 PT disorders related to the musculoskeletal system including musculoskeletal
 PT cancers and also for testing and detection e.g. diagnosis.
 XX
 XX Claim 11; SEQ ID NO 2043; 781pp + Sequence Listing; English.
 PS
 XX The invention relates to novel genes (AAL34669-AAL37666) and proteins
 CC (ABB03087-ABB04109) associated with the musculoskeletal system useful for
 CC preventing, treating or ameliorating medical conditions e.g. by protein
 CC or gene therapy. The genes are isolated from a range of human tissues
 CC disclosed in the specification. The nucleic acids, proteins, antibodies

CC and (ant)agonists are useful in the diagnosis, treatment and prevention
 CC of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the
 CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
 CC lung, or urogenital; (b) immune disorders e.g. Addison's disease,
 CC allergies, autoimmune haemolytic anaemia, multiple sclerosis, rheumatoid
 CC diabetes mellitus, Crohn's disease, autoimmune thyroiditis,
 CC arthritis and ulcerative colitis; (c) cardiovascular disorders such as
 CC myocardial ischaemias; (d) wound healing; (e) neurological disorders such as
 CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
 CC bacterial, fungal and parasitic infections. Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIFO at
 CC ftp.wifo.int/pub/published_pct_sequences
 XX
 XX Sequence 74 AA;
 SQ
 Query Match 53.4%; Score 39; DB 4; Length 74;
 Best Local Similarity 63.6%; Pred. No. 87;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 3 RCGGDSKVOE 13
 Db 23 RCGGDSRSQE 33
 RESULT 39
 ABUI3390
 ID ABUI3390 standard; protein; 74 AA.
 XX
 AC ABUI3390;
 XX
 DT 26-FEB-2003 (first entry)
 XX
 DE Novel human musculoskeletal system antigen #1010.
 XX
 KW Musculoskeletal system antigen; cancer; metastasis; re-vascularisation;
 KW thrombosis; arteriosclerosis; mineral content; cardiovascular condition;
 KW wound; injury; burn; angiogenesis; ulcer; post-operative tissue repair;
 KW limb regeneration; neuronal growth; neurodegenerative disorder;
 KW Alzheimer's disease; Parkinson's disease; AIDS-related complex;
 KW chondrocyte growth; bone regeneration; periodontal regeneration;
 KW tissue transport; bone graft; skin aging; keratinocyte growth; hair loss;
 KW melanocyte growth; cell proliferation; cell growth; organ transplant;
 KW cell differentiation; body height; weight; hair colour; eye colour; skin;
 KW percentage of adipose tissue; pigmentation; cosmetic surgery; metabolism;
 KW biorhythm; cardiac rhythm; depression; tendency for violence; pain;
 KW reproductive capability; hormone level; endocrine level; appetite;
 KW libido; memory; stress; storage capability; fat content; lipid content;
 KW protein content; carbohydrate content; vitamin content; cofactor content;
 KW nutritional component.
 OS Homo sapiens.
 XX
 XX US2002147140-A1.
 PN
 XX
 XX 10-OCT-2002.
 PD
 XX
 XX 17-JAN-2001; 2001US-00764877.
 PF
 XX 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216800P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 08-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.

XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.

XX Rosen CA, Ruben SM, Barash SC;
XX WPI; 2003-128199/12.

XX Isolated nucleic acid molecules encoding musculoskeletal system
PT associated polypeptides, useful for detecting disorders, e.g. cancer.

XX Disclosure; SEQ ID NO 2043; 321pp; English.

CC The invention describes an isolated nucleic acid molecule comprising a
CC sequence encoding musculoskeletal system associated polypeptides useful
CC for detecting disorders, e.g., cancer or cancer metastases, in animals or
CC humans. The nucleic acid; stimulates re-vascularisation of ischaemic
CC tissues associated with conditions such as thrombosis, arteriosclerosis,
CC and other cardiovascular conditions; treats wounds due to injuries,
CC burns, post-operative tissue repair, and ulcers; stimulates angiogenesis
CC and limb regeneration; stimulates neuronal growth; can treat and prevent
CC neuronal damage occurring in certain disorders or neurodegenerative
CC conditions, such as, Alzheimer's disease, Parkinson's disease, and AIDS-
CC related complex; stimulates chondrocyte growth, thus they can be used to
CC enhance bone and periodontal regeneration and aid in tissue transports or
CC bone grafts; prevents skin aging due to sunburn by stimulating
CC keratinocyte growth; prevents hair loss, since FGF family members
CC activate hair-forming cells and promotes melanocyte growth; stimulates
CC growth and differentiation of hematopoietic cells and bone marrow cells
CC when used in combination with other cytokines; maintains organs before
CC transplantation or for supporting cell culture of primary tissues;
CC induces tissue of mesodermal origin to differentiate in early embryos;
CC increases or decreases the differentiation or proliferation of embryonic
CC stem cells, besides, haematopoietic lineage; modulates mammalian
CC characteristics, such as, body height, weight, hair colour, eye colour,
CC skin, percentage of adipose tissue, pigmentation, size, and shape (e.g.,

CC cosmetic surgery); modulates mammalian metabolism; changes mammal's metal
CC state or physical state by influencing biorhythms, cardiac rhythms,
CC depression, tendency for violence, tolerance for pain, reproductive
CC capabilities, hormonal or endocrine levels, appetite, libido, memory, or
CC stress; increases or decreases storage capabilities, fat content, lipid,
CC protein, carbohydrate, vitamins, minerals, cofactors or other nutritional
CC components. This is the amino acid sequence of a novel human
CC musculoskeletal system antigen. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the US patent office at
XX ftp.seqdata.uspto.gov/sequence.html?DocID=20020147140

XX Sequence 74 AA;

Query Match 53.4%; Score 39; DB 6; Length 74;
Best Local Similarity 63.6%; Pred. No. 87;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 RCRGDDSKVQE 13
Db 23 RCGRDSRSQ 33
||:|:|:|:|

RESULT 40

ADJ29416
ID ADJ29416 standard; protein; 74 AA.

XX ADJ29416;

XX 20-MAY-2004 (first entry)

DE Human musculoskeletal system-associated protein - SEQ ID 2043.

XX musculoskeletal system; cytostatic; osteopathic; cancer; osteoporosis;
XX gene therapy; vaccine; human.

OS Homo sapiens.

XX US2004009488-A1.

XX 15-JAN-2004.

XX 13-SEP-2002; 2002US-00242515.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225477P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

```
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-02331968P.
PR 14-SEP-2000; 2000US-02332397P.
PR 14-SEP-2000; 2000US-02332398P.
PR 14-SEP-2000; 2000US-02332399P.
PR 14-SEP-2000; 2000US-0233400P.
PR 14-SEP-2000; 2000US-0233401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234998P.
PR 25-SEP-2000; 2000US-0234999P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241807P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244826P.
PR 08-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 17-NOV-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764877.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Barash SC;
PI WPI; 2004-090458/09.
DR N-PSDB; ADJ28393.
XX
PT New nucleic acid molecule, useful for preparing a medicament for
PT preventing, treating or ameliorating a medical condition e.g., cancer of
PT musculoskeletal tissues or osteoporosis.
XX Claim 11; SEQ ID NO 2043; 289pp; English.
XX The invention relates to a novel isolated musculoskeletal system-
XX associated nucleic acid molecule. The nucleic acid of the invention
XX demonstrates cytostatic and osteopathic activities and may be useful for
XX preparing a medicament for preventing, treating or ameliorating a medical
XX condition such as cancer of the musculoskeletal tissues or osteoporosis,
XX possibly via gene therapy or vaccine production. The current sequence is
XX that of the human musculoskeletal system-associated polypeptide of the
XX invention. The current sequence is not shown within the specification per
XX se but is available on the USPTO web-site
XX http://seqdata.uspto.gov/sequence.html?DocID=20040009488.
SQ Sequence 74 AA;
Query Match 53.4%; Score 39; DB 8; Length 74;
Best Local Similarity 63.6%; Pred No. 87;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 3 RCRGDSKVQE 13
|||: |||: |||
Db 23 RCQGRDSRSQE 33
RESULT 41
ADE61256
```


CC (M1) are useful in personal hygiene products. The protein esters of (M1)
 CC are useful in shower and bath preparations and in shampoos and body
 CC cleansers. The protein-fatty acid condensates produces by (M1) form a
 CC protective layer on the skin, which reduces the excessive attack of
 CC surfactants on the upper layer of the skin, their strong degreasing
 CC effect and the direct interaction of anionic surfactants with the skin.
 CC The carbohydrate esters produced by (M1) are useful in oral-care
 CC products, as antibiotics, antitumorals, fungicides and insecticides. The
 CC esters produced by (M1) are useful in cosmetics including essential oil
 CC emulsions, paraffin oil emulsions, pharmaceutical preparations including
 CC drug emulsions, ointment bases, agricultural including cleaners for fruit
 CC and vegetables, crop protection including insecticides and herbicides,
 CC and food industry including bread, cakes, chocolate, margarine, custard
 CC powder, and in drinks industry. The present sequence is one such lipid
 CC acyltransferase, used in the method of the invention.

XX SQ Sequence 305 AA;

Query Match 53.4%; Score 39; DB 8; Length 305;
 Best Local Similarity 63.6%; Pred. No. 3.7e+02;
 Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 RCRGDDSKVQE 13
 |||||
 Db 258 RCRSDDGAVHE 268

RESULT 43

ADRA41585
 ID ADRA41585 standard; protein; 305 AA.

AC ADRA41585;

XX 21-OCT-2004 (first entry)

XX DE Lipid acyltransferase, SEQ ID 28.

XX KW Antiarteriosclerotic; Cardiant; Antilipemic; cholesterol; lipoprotein;
 KW emulsifier; foodstuff; lipid acyltransferase; enzyme;
 KW blood serum cholesterol; low density lipoprotein; atherosclerosis;
 KW heart disease.

XX OS Streptomyces coelicolor A3(2).

XX PN WO2004064537-A2.

XX PD 05-AUG-2004.

XX PF 15-JAN-2004; 2004WO-IB000655.

XX PR 17-JAN-2003; 2003GB-00001117.

XX PR 17-JAN-2003; 2003GB-00001118.

XX PR 17-JAN-2003; 2003GB-00001119.

XX PR 17-JAN-2003; 2003GB-00001120.

XX PR 17-JAN-2003; 2003GB-00001121.

XX PR 17-JAN-2003; 2003GB-00001122.

XX PR 23-JUL-2003; 2003US-0489441P.

XX PR 24-DEC-2003; 2003GB-00030016.

XX PA (DANI-) DANISCO AS.

XX PI De Kreij A, Madrid SM, Mikkelsen JD, Soe JB;

XX DR WPI; 2004-593271/57.

XX DR N-PSDB; ADRA41586.

XX PT In situ production of emulsifier in foodstuffs, involves addition of
 PT lipid acyltransferase to foodstuff.

XX PS Claim 13; Fig 24; 257pp; English.

XX CC The present invention relates to a method (M1) for the in situ production
 CC of an emulsifier in a foodstuff, wherein a lipid acyltransferase (I) is

CC added to the foodstuff. (I) enables in situ emulsifier production,
 CC without increasing or without substantially increasing the free fatty
 CC acid in the foodstuffs. (I) also produces foodstuffs having improved
 CC appearance, stability, thermal stability, taste, softness, resilience and
 CC emulsification. (I) comprises the motif ADR41613. (M1) or (I) is useful
 CC for in situ production of an emulsifier in a foodstuff. (I) is useful for
 CC preparing a foodstuff (e.g., egg-based products, salad dressings sauces,
 CC ice creams, egg powder, modified egg yolk or bread) which in turn is
 CC useful for reducing blood serum cholesterol and/or to reduce low density
 CC lipoprotein, or for treating and/or preventing atherosclerosis and/or
 CC heart disease. (I) is useful for preparing medicament, flavouring agent
 CC and/or texturizer. (I) is useful for producing acylated proteins,
 CC cosmetic composition or food or feed enzyme composition. The present
 CC sequence is one such lipid acyltransferase used to illustrate the
 CC invention.

XX SQ Sequence 305 AA;

Query Match 53.4%; Score 39; DB 8; Length 305;
 Best Local Similarity 63.6%; Pred. No. 3.7e+02;
 Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 RCRGDDSKVQE 13
 |||||
 Db 258 RCRSDDGAVHE 268

RESULT 44

ADN48160
 ID ADN48160 standard; protein; 393 AA.

XX AC ADN48160;

XX DT 01-JUL-2004 (first entry)

XX DE Thermococcus kodakaraensis KOD1 protein sequence SeqID2038.

XX KW gene disruption; gene targeting; marker gene; transformation;
 KW homologous recombination; hyperthermostable archaeobacterium; KOD1;
 KW gene structure; gene function; enzyme activity; medicine; immunology.
 KW forensic science; food; drug inspection; molecular biology; immunology.

XX OS Thermococcus kodakaraensis.

XX PN WO2004022736-A1.

XX PD 18-MAR-2004.

XX PF 29-AUG-2003; 2003WO-IB003597.

XX PR 30-AUG-2002; 2002JP-00319011.

XX PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.

XX PI Imanaka T, Atomi H;

XX DR WPI; 2004-257583/24.

XX PT Method for disrupting targeted gene in genome of organism particularly
 PT thermostable bacterium and with genome chips for analysis, applicable in
 PT studying gene structure and functions.

XX PS Claim 9; SEQ ID NO 2038; 598pp; Japanese.

XX CC This invention relates to a novel method for targeting disruption of an
 CC arbitrary gene in a genome of an organism which comprises providing the
 CC whole sequential data of the genome of such organism, selecting at least
 CC 1 arbitrary region in the sequence, providing a vector that contains a
 CC sequence homologous with the selected region and a marker gene,
 CC transformation, and homologous recombination. The genome is preferably
 CC the genome of a hyperthermostable archaeobacterium, particularly
 CC Thermococcus kodakaraensis KOD1. The method is for targeting the
 CC disruption of a gene in the genome of an organism, which is applicable in

CC studying gene structure and functions as well as enzyme activities of
CC encoded proteins and useful in medicine, forensic science, food or drug
CC inspection, molecular biology and immunology. With this method, the
CC disruption of a gene at an arbitrary position in a genome can be achieved
CC efficiently and reliably. The present sequence is that of a protein
CC encoded by the genome of *Thermococcus kodakaraensis* which was derived
CC using the method of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 393 AA;
SQ

Query Match 53.4%; Score 39; DB 8; Length 393;
Best Local Similarity 53.8%; Pred. No. 4.8e+02;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 NYRCRGDDSKVQ 13
Db 279 STRAGDDKATQE 291

RESULT 45
ADG33762
ID ADG33762 standard; protein; 461 AA.
XX AC ADG33762;
XX
XX
XX 26-FEB-2004 (first entry)
XX
XX Actinomycetes dual condensation/epimerisation NRPS domain protein ID 21.
XX
XX non-ribosomal peptide synthetase; NRPS; dual condensation; epimerisation;
KW ramoplanin; complestatin; actinomycetes taxon.
XX
XX *Pseudomonas syringae* pv. *syringae*.
XX
XX WO2003089641-A2.
XX
XX 30-OCT-2003.
XX
XX 17-APR-2003; 2003WO-CA000575.
XX
XX 17-APR-2002; 2002US-0372790P.
PR (ECOP-) ECOPIA BIOSCIENCES INC.
XX
XX Farnet CM, Staffa A;
XX
XX WPI; 2003-854123/79.
DR N-PSDB; ADG33763.
XX
XX New dual condensation/epimerization non-ribosomal peptide synthetase
PT domain and encoding polynucleotide, useful for modifying the
PT stereochemistry of synthesized peptides (e.g. ramoplanin or complestatin)
PT in vitro or in vivo.
XX
XX Claim 9; SEQ ID NO 21; 245pp; English.
XX
XX This invention relates to novel domains of non-ribosomal peptide
CC synthetases (NRPSs) that exhibit dual condensation and epimerisation
CC activities. Specifically, these domains allow incorporation of non-
CC proteinogenic substrates (e.g. D-amino acids) into peptide products.
CC Furthermore, they can be used in vivo to modify the stereochemistry of
CC synthesized peptides (e.g. ramoplanin or complestatin) at selected amino
CC acid sites by the addition of non-chiral residues. The present invention
CC describes the identification of isolated polynucleotide NRPS domains in
CC various organisms from the actinomycetes taxon, and the encoded
CC polypeptides thereof, as well as suitable expression vectors. This
CC polypeptide sequence is a dual condensation/ epimerisation NRPS domain
CC protein of the invention.
XX
XX Sequence 461 AA;
SQ

Query Match 53.4%; Score 39; DB 7; Length 461;
Best Local Similarity 50.0%; Pred. No. 5.7e+02;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDDSKVQ 12
Db 341 NYRHQGDNRLLQ 352

RESULT 46
ADS44300
ID ADS44300 standard; protein; 690 AA.
XX AC ADS44300;
XX
XX 02-DEC-2004 (first entry)
XX
XX Bacterial polypeptide #22730.
XX
XX Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; omeosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
XX Bacteria.
XX
XX US2003233675-A1.
XX
XX 18-DEC-2003.
XX
XX 20-FEB-2003; 2003US-00369493.
XX
XX 21-FEB-2002; 2002US-0360039P.
XX
XX (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 22730; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress

CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 690 AA;

SQ Query Match 53.4%; Score 39; DB 8; Length 690;

Best Local Similarity 80.0%; Pred. No. 8.6e+02; Mismatches 0; Indels 0; Gaps 0;
 Matches 8; Conservative 0;

QY 3 RCRGDDSKVQ 12

Db 561 RVRGDDSDVQ 570

RESULT 47

ADD46237
 ID ADD46237 standard; protein; 1269 AA.

XX ADD46237;

XX 29-JAN-2004 (first entry)

XX Rat Protein U49055, SEQ ID NO 11912.

XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 KW Chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX Rattus norvegicus.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

XX 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

XX (GHEO) GEN HOSPITAL CORP.

XX (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX GENBANK; U49055.

XX New composition comprising two or more isolated polypeptides, useful for

PT preparing a medicament for treating pain in an animal.

PS Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the

XX polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more

CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1269 AA;

SQ Query Match 52.7%; Score 38.5; DB 7; Length 1269;

Best Local Similarity 72.7%; Pred. No. 2e+03; Mismatches 1; Indels 1; Gaps 1;
 Matches 8; Conservative 1;

QY 2 YRCRGDD-SKV 11

Db 100 YRCPGDDSKI 110

RESULT 48

ADD46239
 ID ADD46239 standard; protein; 1272 AA.

XX AC ADD46239;

XX 29-JAN-2004 (first entry)

XX Human Protein AB029039, SEQ ID NO 11914.

XX Human; pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; chronic constriction injury; CCI;
 KW spared nerve injury; SNI; Chung.

XX Homo sapiens.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

XX 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

XX (GHEO) GEN HOSPITAL CORP.

XX (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX GENBANK; AB029039.

XX New composition comprising two or more isolated polypeptides, useful for

PT preparing a medicament for treating pain in an animal.

PS Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the

CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a human protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 1272 AA;
 Query Match 52.7%; Score 38.5; DB 7; Length 1272;
 Best Local Similarity 72.7%; Pred. No. 2e+03;
 Matches 8; Conservative 1; Mismatches 1; Indels 1; Gaps 1;
 QY 2 YRCRGDD-SKV 11
 ||| ||| |||
 Db 100 YRCFGDDSKI 110

RESULT 49
 AAU60236
 ID AAU60236 standard; protein; 84 AA.
 XX
 AC AAU60236;
 XX
 DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #21132.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.
 OS
 XX WO200181581-A2.
 PN
 XX
 PD 01-NOV-2001.
 XX
 XX 20-APR-2001; 2001WO-US012865.
 XX
 XX 21-APR-2000; 2000US-0199047P.
 PR
 PR 02-JUN-2000; 2000US-0208841P.
 PR
 PR 07-JUL-2000; 2000US-0216747P.

XX (CORI-) CORIXA CORP.
 XX
 XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'Maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 XX WPI; 2001-616774/71.
 DR
 DR N-PSDB; AAS59608.
 XX
 XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.

XX Example 1; SEQ ID NO 21431; 1069pp; English.
 PS
 XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central

CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 84 AA;
 Query Match 52.1%; Score 38; DB 4; Length 84;
 Best Local Similarity 38.5%; Pred. No. 1.5e+02;
 Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
 QY 1 NYRCRGDDSKVQ 13
 : ||| : : :
 Db 36 HYACRGEGVRIQD 48

RESULT 50
 ABM56755
 ID ABM56755 standard; protein; 84 AA.
 XX
 AC ABM56755;
 XX
 DT 20-OCT-2003 (first entry)
 XX
 DE Propionibacterium acnes predicted ORF-encoded polypeptide #21431.
 XX
 KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;
 KW immunostimulant; immune response; vaccine.

XX Propionibacterium acnes.
 OS
 XX WO2003033515-A1.
 PN
 XX 24-APR-2003.
 PD
 XX 11-OCT-2002; 2002WO-US032727.
 PF
 XX 15-OCT-2001; 2001US-00978825.
 PR
 XX (CORI-) CORIXA CORP.
 XX
 XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
 PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
 PI Barth B, Vallieve-Douglas J;
 XX
 XX WPI; 2003-381789/36.
 DR
 DR N-PSDB; ACF64537.

XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
 PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
 PT or for stimulating an immune response specific for a P. acnes protein.
 XX
 XX Example 1; SEQ ID NO 21431; 1481pp; English.

XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
 CC encoding a Propionibacterium acnes protein. The invention also relates to
 CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
 CC immunogenic fragments of P. acnes polypeptides. The invention
 CC additionally encompasses expression vectors and host cells comprising a
 CC polynucleotide of the invention; antibodies against polypeptides of the
 CC polynucleotide of the invention; fusion proteins comprising a polypeptide of the invention; a
 CC method for stimulating an immune response specific for a P. acnes
 CC polypeptide and an isolated T cell population comprising T cells prepared

CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX

SQ Sequence 84 AA;

Query Match 52.1%; Score 38; DB 6; Length 84;
Best Local Similarity 38.5%; Pred. No. 1.5e+02;
Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 NYRCRGDSDSKVQE 13
:| |||: :|:
Db 36 HYACRGEGVRIQD 48

Search completed: September 7, 2005, 19:55:12
Job time : 109.721 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2005, 19:49:10 ; Search time 9.97674 Seconds
(without alignments)
125.373 Million cell updates/sec

Title: US-10-812-238B-23

Perfect score: 73

Sequence: 1 NYRCRDDSKVQE 13

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database :

1: PIR 79:*

2: PIR1:*

3: PIR2:*

4: PIR3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	60.3	504	2 C64398	hypothetical prote
2	44	60.3	705	2 A70669	probable acid-CoA
3	42	57.5	513	2 B87484	anthranilate synth
4	40	54.8	150	2 G83840	ferric uptake regu
5	40	54.8	179	2 T23521	hypothetical prote
6	40	54.8	303	2 T29321	hypothetical prote
7	40	54.8	400	1 A39130	nicotinate phospho
8	40	54.8	400	1 JQ0756	nicotinate phospho
9	40	54.8	400	2 AB0617	nicotinate phospho
10	40	54.8	400	2 F90755	nicotinate phospho
11	40	54.8	400	2 D85619	nicotinate phospho
12	40	54.8	577	2 T18646	hypothetical prote
13	40	54.8	721	2 H82528	L-ascorbate oxidas
14	39	53.4	280	2 T43011	suppressor protein
15	39	53.4	325	2 S66668	hydrogen peroxide-
16	39	53.4	507	2 AD3382	ABC transporter-as
17	39	53.4	662	2 T41442	omnipotent nonsens
18	39	53.4	662	2 T51948	omnipotent nonsens
19	39	53.4	9376	2 T14593	synergomycin synth
20	38.5	52.7	1268	2 T31420	C-terminal domain-
21	38	52.1	172	2 F81439	translation initia
22	38	52.1	389	2 T27085	hypothetical prote
23	38	52.1	1071	2 B85343	hypothetical prote
24	38	52.1	1267	1 MWXR31	lambda 3 protein -
25	38	52.1	1267	1 MWXR32	lambda 3 protein -
26	38	52.1	1267	1 MWXR33	lambda 3 protein -
27	38	52.1	2212	2 T28157	erythrocyte membra
28	38	52.1	2647	2 T28161	hypothetical prote
29	38	52.1	3228	2 T21381	hypothetical prote

30	37	50.7	185	2 H87107	ribosome recycling
31	37	50.7	258	2 T21959	hypothetical prote
32	37	50.7	324	2 H86244	lysophospholipase
33	37	50.7	410	2 T21960	hypothetical prote
34	37	50.7	430	2 S75313	serine-tRNA ligase
35	37	50.7	489	2 C87480	conserved hypotet
36	37	50.7	518	2 T19133	hypothetical prote
37	37	50.7	564	2 T19132	hypothetical prote
38	37	50.7	801	2 A89862	Na+/H+ antiporter
39	37	50.7	873	2 T19131	hypothetical prote
40	37	50.7	988	2 F86316	protein T10022.13
41	37	50.7	1030	2 F96763	hypothetical prote
42	36	49.3	138	2 C90460	hypothetical prote
43	36	49.3	180	2 AC3270	hypothetical expor
44	36	49.3	248	2 F64908	probable dehydroge
45	36	49.3	248	2 D90897	probable oxidoredu
46	36	49.3	248	2 C85720	probable oxidoredu
47	36	49.3	329	2 T20546	hypothetical prote
48	36	49.3	347	2 T35013	probable membrane
49	36	49.3	411	2 H74494	hypothetical prote
50	36	49.3	435	2 A46206	voltage-gated sodi
51	36	49.3	461	1 E69360	4-hydroxyphenylac
52	36	49.3	470	2 E72376	hypothetical prote
53	36	49.3	644	2 B72559	probable arginyl-t
54	36	49.3	707	2 F86925	probable acyl-CoA
55	36	49.3	752	1 P11VBC	RNA-directed RNA p
56	36	49.3	752	1 P11VBL	RNA-directed RNA p
57	36	49.3	752	1 P11VBW	RNA-directed RNA p
58	36	49.3	822	2 S77112	DNA mismatch repa
59	36	49.3	1812	2 I49350	breast/ovarian can
60	35	47.9	187	2 A10520	AmpD protein (anhy
61	35	47.9	200	2 T02826	ribosomal protein
62	35	47.9	269	2 A46506	leukocyte activati
63	35	47.9	274	2 T32736	hypothetical prote
64	35	47.9	342	2 T45577	protein kinase-lik
65	35	47.9	346	2 T45085	hypothetical prote
66	35	47.9	347	2 B72596	hypothetical prote
67	35	47.9	392	2 A86252	hypothetical prote
68	35	47.9	422	2 S34199	polygalacturonase
69	35	47.9	435	2 T32195	hypothetical prote
70	35	47.9	453	2 T24127	probable chitinase
71	35	47.9	471	2 S09110	hypothetical prote
72	35	47.9	482	2 H71400	hypothetical prote
73	35	47.9	502	2 S38139	LAS1 protein - yea
74	35	47.9	656	2 T10664	serine/threonine-s
75	35	47.9	662	2 I37892	IL12 receptor comp
76	35	47.9	1035	2 T42093	phospholipase D (E
77	35	47.9	1036	2 T13732	phospholipase D (E
78	35	47.9	1037	2 T13943	phospholipase D (E
79	35	47.9	1074	2 T17203	phospholipase (EC
80	35	47.9	1074	2 T13725	phospholipase D (E
81	35	47.9	1075	2 T46635	phospholipase D (E
82	35	47.9	1323	2 T37533	coronin-like prote
83	35	47.9	1390	2 S51364	sperm tail-specifi
84	35	47.9	2153	2 T14893	scavenger receptor
85	34.5	47.3	325	2 B81364	probable periplasm
86	34	46.6	112	2 H70232	hypothetical prote
87	34	46.6	116	2 H75591	hypothetical prote
88	34	46.6	122	2 T04118	mitochondrial proc
89	34	46.6	137	2 T33210	hypothetical prote
90	34	46.6	225	2 T20325	hypothetical prote
91	34	46.6	238	2 T11718	hypothetical prote
92	34	46.6	255	2 G64320	hypothetical prote
93	34	46.6	259	2 E84427	hypothetical prote
94	34	46.6	270	2 C96731	unknown protein F5
95	34	46.6	273	2 T10610	hypothetical prote
96	34	46.6	285	2 S36903	Fc gamma (IgG) rec
97	34	46.6	296	2 A95408	hypothetical prote
98	34	46.6	302	2 S65021	chitinase (EC 3.2.
99	34	46.6	308	2 D86742	thioredoxin-disulf
100	34	46.6	311	2 F84030	porphobilinogen de
101	34	46.6	313	1 I8EC	hydroxymethylbilan
102	34	46.6	324	2 S56694	chitinase (EC 3.2.

103	34	46.6	327	1	JQ0965	chitinase (EC 3.2.	176	33	45.2	370	2	F84550	hypothetical prote
104	34	46.6	341	2	H95930	hypothetical prote	177	33	45.2	372	2	A34261	alanine dehydrogen
105	34	46.6	344	2	T05437	hypothetical prote	178	33	45.2	375	2	AF1720	N-acetylglucosaminidase
106	34	46.6	348	2	H85256	hypothetical prote	179	33	45.2	387	2	JL0101	methane monooxygenase
107	34	46.6	358	2	T32238	hypothetical prote	180	33	45.2	389	1	A48329	histidinol-phosphatase
108	34	46.6	375	2	B96551	hypothetical prote	181	33	45.2	393	2	A75079	hypothetical prote
109	34	46.6	375	2	AC1350	hypothetical prote	182	33	45.2	408	2	S76678	hypothetical prote
110	34	46.6	393	2	B86189	N-acetylglucosaminidase	183	33	45.2	421	2	AE3467	glutamate dehydrogenase
111	34	46.6	407	2	C36786	hypothetical prote	184	33	45.2	428	2	S58735	homocitrate synthase
112	34	46.6	413	2	F96743	hypothetical prote	185	33	45.2	435	2	A49171	translation elongation factor
113	34	46.6	505	2	A55923	KOR4 protein - yea	186	33	45.2	438	2	E70883	probable proteinase
114	34	46.6	521	2	T18896	glypican 1 precursor	187	33	45.2	456	2	S11665	translation elongation factor
115	34	46.6	559	2	S04531	thrombospondin-rel	188	33	45.2	460	2	JC4214	translation elongation factor
116	34	46.6	572	1	VGNZBA	cell fusion glycoprotein	189	33	45.2	484	2	E64432	spore coat polysaccharide
117	34	46.6	596	2	B41627	cell fusion glycoprotein	190	33	45.2	503	2	T46852	anthranilate synthase
118	34	46.6	692	2	T32980	hypothetical prote	191	33	45.2	514	2	T00934	probable cytochrome
119	34	46.6	710	1	IS1283	hepatocyte growth factor	192	33	45.2	524	2	A32617	probable cytochrome
120	34	46.6	731	2	T29129	probable ATP/GTP b	193	33	45.2	527	2	T41856	ACMNPV orf119 - Bo
121	34	46.6	732	1	HYHUK	Kell blood group p	194	33	45.2	531	2	T51922	hypothetical prote
122	34	46.6	757	1	A60008	RNA-directed RNA p	195	33	45.2	532	2	S18448	variant surface glycoprotein
123	34	46.6	757	1	B60011	RNA-directed RNA p	196	33	45.2	552	2	S48328	hypothetical prote
124	34	46.6	757	1	PL1V61	RNA-directed RNA p	197	33	45.2	569	2	D85354	hypothetical prote
125	34	46.6	757	1	PL1V33	RNA-directed RNA p	198	33	45.2	573	2	H86313	protein F2H15.10 l
126	34	46.6	757	1	PL1V34	RNA-directed RNA p	199	33	45.2	574	1	VGNZB2	cell fusion glycoprotein
127	34	46.6	757	1	PL1V68	RNA-directed RNA p	200	33	45.2	574	1	VGNZB2	cell fusion glycoprotein
128	34	46.6	757	2	S06212	RNA-directed RNA p	201	33	45.2	574	1	VGNZB2	cell fusion glycoprotein
129	34	46.6	783	2	A41627	RNA-directed RNA p	202	33	45.2	574	1	VGNZB2	cell fusion glycoprotein
130	34	46.6	808	2	H64474	hypothetical prote	203	33	45.2	574	2	B28929	cell fusion glycoprotein
131	34	46.6	850	2	S54553	hypothetical prote	204	33	45.2	574	2	S37254	cell fusion glycoprotein
132	34	46.6	1036	2	T18530	phospholipase D (8	205	33	45.2	587	2	C71433	probable membrane
133	34	46.6	1116	2	T38073	serine/threonine-p	206	33	45.2	589	2	F70033	glucan 1,4-alpha-mannanase
134	34	46.6	1494	2	T13798	hypothetical prote	207	33	45.2	600	2	E87665	arginyl-tRNA synthetase
135	34	46.6	1750	2	H64403	ribonucleoside-tri	208	33	45.2	662	2	T18233	probable transcrip
136	34	46.6	1763	2	T19183	hypothetical prote	209	33	45.2	675	2	S74399	sensory transducti
137	34	46.6	1769	2	T13184	hypothetical prote	210	33	45.2	680	2	G82526	ATP-dependent DNA
138	33-5	45.9	348	2	F87076	probable reductase	211	33	45.2	718	1	A69084	DNA topoisomerase
139	33-5	45.9	368	2	G69142	GDP-D-mannose dehy	212	33	45.2	725	2	JE0100	neural cell adhesi
140	33-5	45.9	370	2	T22082	hypothetical prote	213	33	45.2	730	2	JH0798	fasciclin IV precu
141	33	45.2	60	2	A54369	fibrogen recepto	214	33	45.2	789	2	T51310	RNA helicase RH28
142	33	45.2	112	2	B86263	FLK23.19 protein	215	33	45.2	842	2	C83177	probable phosphotr
143	33	45.2	152	2	AI0448	probable ribonucle	216	33	45.2	860	2	T16892	hypothetical prote
144	33	45.2	161	2	AC0883	conserved hypotet	217	33	45.2	864	2	T30441	probable capsid-as
145	33	45.2	162	2	F82084	probable 2-demethy	218	33	45.2	938	2	A56731	chromatin assembly
146	33	45.2	167	2	B84536	hypothetical prote	219	33	45.2	955	2	T18435	hypothetical prote
147	33	45.2	169	2	C70207	outer membrane pro	220	33	45.2	969	2	H69425	modification methy
148	33	45.2	170	2	AC2917	invasion protein A	221	33	45.2	996	2	T25512	hypothetical prote
149	33	45.2	175	2	T18773	hypothetical prote	222	33	45.2	1034	2	A36108	integrin alpha-v c
150	33	45.2	184	2	T13186	hypothetical prote	223	33	45.2	1041	2	PQ0442	polyprotein - barl
151	33	45.2	191	2	AG2916	conserved hypotet	224	33	45.2	1044	2	T10050	integrin alpha-v c
152	33	45.2	194	2	F97691	hypothetical prote	225	33	45.2	1045	2	S60571	integrin alpha-5 c
153	33	45.2	195	2	S39667	ribonuclease (EC 3	226	33	45.2	1048	2	A27421	integrin alpha-5 c
154	33	45.2	198	2	S06176	cytotoxic T-lympho	227	33	45.2	1092	1	JN0635	neural cell adhesi
155	33	45.2	205	2	S34833	stylar protein (al	228	33	45.2	1174	2	S28976	DNA-directed RNA p
156	33	45.2	210	2	T18778	hypothetical prote	229	33	45.2	1202	2	PQ0440	polyprotein - barl
157	33	45.2	214	2	S46476	cysteine proteinase	230	33	45.2	1231	1	NBHUH	complement factor
158	33	45.2	219	2	T10520	ribonuclease (EC 3	231	33	45.2	1289	2	AB2217	hypothetical prote
159	33	45.2	221	2	T04238	hypothetical prote	232	33	45.2	1436	2	A46496	antigen Wc1.1 prec
160	33	45.2	226	2	AD2198	hypothetical prote	233	33	45.2	1470	2	S45323	genome polyprotein
161	33	45.2	235	2	B97691	phbP protein (AF02	234	33	45.2	2043	2	T18524	scavenger receptor
162	33	45.2	242	2	A36148	flbN protein precu	235	33	45.2	2051	2	T30938	receptor tyrosine
163	33	45.2	244	2	A87505	flagellar L-ring p	236	33	45.2	2410	1	JQ1948	genome polyprotein
164	33	45.2	248	2	A33412	cytotoxic T-lympho	237	33	45.2	2412	1	JQ1537	LDL-receptor-relat
165	33	45.2	273	2	D81878	probable oxidoredu	238	33	45.2	4753	1	A47437	LDL-receptor-relat
166	33	45.2	324	2	S20981	chitinase (EC 3.2.	239	32.5	44.5	186	2	AE3624	hypothetical prote
167	33	45.2	328	2	AG2987	hypothetical prote	240	32.5	44.5	435	2	C59194	probable sulfatase
168	33	45.2	329	2	S08627	chitinase (EC 3.2.	241	32.5	44.5	571	2	AI0506	hypothetical prote
169	33	45.2	333	2	A98296	hypothetical prote	242	32	43.8	46	2	D84334	DNA-binding protei
170	33	45.2	336	2	G82228	forminoglutamate	243	32	43.8	66	1	DNVBP	DNA-binding protei
171	33	45.2	336	2	T31249	hypothetical prote	244	32	43.8	74	1	DNVBPAS	preoptic area regu
172	33	45.2	350	2	D83102	xenobiotic reducta	245	32	43.8	75	2	B36119	hypothetical prote
173	33	45.2	352	2	C96643	hypothetical prote	246	32	43.8	97	2	A82158	hypothetical prote
174	33	45.2	352	2	C84603	probable pectinase	247	32	43.8	98	2	F70769	hypothetical prote
175	33	45.2	362	2	T49078	probable serine/th	248	32	43.8	111	2	C71401	hypothetical prote

249	32	43.8	116	2	H69023	conserved hypothet	322	32	43.8	486	2	T21566	hypothetical prote
250	32	43.8	129	2	AF0527	conserved hypothet	323	32	43.8	489	2	C84725	probable protein k
251	32	43.8	142	2	JH0586	ribosomal protein	324	32	43.8	491	2	E87452	ankyrin-related pr
252	32	43.8	145	2	A83010	conserved hypothet	325	32	43.8	504	2	F82253	amidophosphoribos
253	32	43.8	150	2	T25581	hypothetical prote	326	32	43.8	505	2	T19106	probable serine ca
254	32	43.8	152	2	C83896	potassium channel	327	32	43.8	509	2	T02864	probable Zn finger
255	32	43.8	154	2	A60398	replication protei	328	32	43.8	516	2	D26682	protein FLE22.18 (
256	32	43.8	158	2	AB0426	transcription elon	329	32	43.8	516	2	T06716	hypothetical prote
257	32	43.8	161	1	TPH0CC	troponin C, cardia	330	32	43.8	522	2	S33029	hypothetical prote
258	32	43.8	161	1	TPBOCC	troponin C, cardia	331	32	43.8	531	2	S07881	hypothetical prote
259	32	43.8	161	1	TPRBCW	troponin C - cardia	332	32	43.8	534	2	T19944	hypothetical prote
260	32	43.8	161	1	S07450	troponin C - quail	333	32	43.8	536	2	T20736	hypothetical prote
261	32	43.8	161	2	A27204	troponin C, cardia	334	32	43.8	539	2	AS4294	cell division cont
262	32	43.8	161	2	JU0035	troponin C, cardia	335	32	43.8	551	2	S58238	DHR38 protein - fr
263	32	43.8	161	2	JW0064	slow cardiac tropo	336	32	43.8	554	2	B85072	hypothetical prote
264	32	43.8	161	2	A32620	troponin C, cardia	337	32	43.8	562	2	AF0852	secretory protein i
265	32	43.8	171	1	RWHUD1	T-cell surface gly	338	32	43.8	563	2	SS4420	invasion protein i
266	32	43.8	176	2	G70138	purine-binding che	339	32	43.8	576	2	T41587	probable carbon ca
267	32	43.8	178	2	AB3279	invasion protein A	340	32	43.8	581	2	F84657	hypothetical prote
268	32	43.8	183	2	AF0304	probable exported	341	32	43.8	591	2	T02402	beta-glucosidase h
269	32	43.8	190	2	T46053	ADP-ribosylation f	342	32	43.8	606	2	T16449	hypothetical prote
270	32	43.8	204	2	E84173	hypothetical prote	343	32	43.8	607	2	S77092	hypothetical prote
271	32	43.8	221	2	T26921	hypothetical prote	344	32	43.8	632	2	T00325	hypothetical prote
272	32	43.8	226	2	T14675	hypothetical prote	345	32	43.8	646	2	S30180	phosphoenolpyruvat
273	32	43.8	229	2	I46614	nerve growth facto	346	32	43.8	660	2	T20569	hypothetical prote
274	32	43.8	229	2	B95147	liC protein (impo	347	32	43.8	664	2	E70770	probable atp-depen
275	32	43.8	229	2	H98014	hypothetical prote	348	32	43.8	710	1	S70965	serine/threonine-s
276	32	43.8	232	2	T14939	hypothetical prote	349	32	43.8	729	2	T45951	hypothetical prote
277	32	43.8	234	2	I56140	NK and T lymphocy	350	32	43.8	729	2	T23972	hypothetical prote
278	32	43.8	241	2	JL0097	nerve growth facto	351	32	43.8	736	2	D86271	protein F16A14.2 (
279	32	43.8	245	2	I56570	beta-nerve growth	352	32	43.8	738	2	A40096	platelet-endotheli
280	32	43.8	245	2	D82880	conserved hypothet	353	32	43.8	767	2	AE3370	sensory transducti
281	32	43.8	248	1	S01007	granzyme F (EC 3.4	354	32	43.8	775	2	S35543	DNA-directed DNA p
282	32	43.8	253	2	S53760	triose-phosphate i	355	32	43.8	862	1	S56766	replication licens
283	32	43.8	253	2	A05283	spectrin alpha cha	356	32	43.8	1027	2	S37711	kinasin heavy chai
284	32	43.8	259	2	T01008	ribonuclease (EC 3	357	32	43.8	1032	2	I38510	neuronal kinesin h
285	32	43.8	286	1	NGHUBM	nerve growth facto	358	32	43.8	1048	2	C86189	protein T2SN20.11
286	32	43.8	287	2	D81295	cytochrome-c oxida	359	32	43.8	1094	2	CS9434	KIAA1688 protein (
287	32	43.8	289	2	T44599	oligopeptide trans	360	32	43.8	1095	2	T00329	hypothetical prote
288	32	43.8	303	1	NGRTBA	nerve growth facto	361	32	43.8	1096	2	T16875	hypothetical prote
289	32	43.8	305	2	E84261	glycosyl transfera	362	32	43.8	1224	2	A25884	DNA-directed RNA p
290	32	43.8	307	1	NGMSMG	nerve growth facto	363	32	43.8	1263	2	T15496	hypothetical prote
291	32	43.8	307	1	AF1590	hypothetical prote	364	32	43.8	1283	2	T13799	neurexin IV - frui
292	32	43.8	320	2	T02285	hypothetical prote	365	32	43.8	1331	2	S05011	calcium channel al
293	32	43.8	321	2	S58686	hypothetical prote	366	32	43.8	1331	2	S05011	calcium channel al
294	32	43.8	345	2	T46962	8-oxoguanine DNA-g	367	32	43.8	1429	2	JN0609	nitric-oxide synth
295	32	43.8	350	2	T25451	transforming growt	368	32	43.8	1433	2	G01946	nitric-oxide synth
296	32	43.8	354	2	T41377	rna binding protei	369	32	43.8	1451	1	JQ1719	E2 glycoprotein pr
297	32	43.8	357	2	AG1072	conserved hypothet	370	32	43.8	1453	2	S41453	spike protein - ca
298	32	43.8	363	2	G95937	probable murein pep	371	32	43.8	1520	2	T30820	carbamoyl-phosphat
299	32	43.8	368	2	S62412	2-dehydro-3-deoxy-	372	32	43.8	1589	1	RGBYCS	cell division cont
300	32	43.8	383	2	A96704	hypothetical prote	373	32	43.8	1984	2	A44396	p-type cation tran
301	32	43.8	385	2	G84051	hypothetical prote	374	32	43.8	2042	2	T18399	variant-specific s
302	32	43.8	387	1	A60345	protein-tyrosine-p	375	32	43.8	2166	2	S11339	calcium channel pr
303	32	43.8	394	2	S77272	hypothetical prote	376	32	43.8	2171	2	S05054	calcium channel al
304	32	43.8	415	1	A33899	protein-tyrosine-p	377	32	43.8	2204	1	RRNZNV	genome polypeptid
305	32	43.8	415	2	T12974	hypothetical prote	378	32	43.8	2274	2	T30258	adenomatous polypo
306	32	43.8	427	2	G02034	hypothetical prote	379	32	43.8	2657	2	T18497	hypothetical prote
307	32	43.8	428	2	S36956	glycolipid 2-alpha	380	32	43.8	2692	2	T23768	hypothetical prote
308	32	43.8	429	2	S23581	lamb protein precu	381	32	43.8	2911	2	T20566	hypothetical prote
309	32	43.8	429	2	S59773	26S proteasome reg	382	32	43.8	4687	1	A39638	plectin - rat
310	32	43.8	444	2	G01924	KIR (cl-2) NK rece	383	31.5	43.2	189	2	G96514	hypothetical prote
311	32	43.8	444	2	G01925	KIR (cl-11) NK rec	384	31.5	43.2	205	2	F87560	ribosomal protein
312	32	43.8	447	2	C82958	probable two-compo	385	31.5	43.2	473	1	A38874	protein-tyrosine-p
313	32	43.8	449	2	F83627	hypothetical prote	386	31.5	43.2	645	2	G72256	hydrogenase (EC 1.
314	32	43.8	455	2	G01923	KIR (cl-5) NK rece	387	31.5	43.2	782	2	T43277	host cell factor 1
315	32	43.8	457	2	T48036	hypothetical prote	388	31.5	43.2	888	2	H88085	protein T11F1.8 (i
316	32	43.8	458	2	S53936	hypothetical prote	389	31.5	43.2	1016	2	T05066	hypothetical prote
317	32	43.8	462	2	H95914	hypothetical membr	390	31.5	43.2	1367	2	H82874	conserved hypothet
318	32	43.8	466	2	H69153	asparagine synthas	391	31.5	43.2	26926	1	I38344	titin, cardiac mus
319	32	43.8	474	1	OMHUIB	alpha-1-B-glycopro	392	31	42.5	77	2	A38345	tachyplestin I prec
320	32	43.8	478	2	JQ2034	RNA-directed RNA p	393	32	42.5	86	2	S38920	hypothetical prote
321	32	43.8	486	2	A83284	probable oxidoredu	394	31	42.5	120	1	PSRF3U	phospholipase A2 (

395	31	42.5	125	2	T32471	hypothetical prote	468	31	42.5	367	1	S02193	cellular tumor ant
396	31	42.5	129	2	AF4902	hypothetical prote	469	31	42.5	369	2	JQ2278	hydroxymethylbilan
397	31	42.5	134	2	AF3482	hypothetical prote	470	31	42.5	374	2	AG3337	transcription regu
398	31	42.5	135	2	I49275	protein kinase Sry	471	31	42.5	377	2	B72275	probable aspartate
399	31	42.5	137	2	E86251	protein F25C20.8 [472	31	42.5	384	2	A11843	hypothetical prote
400	31	42.5	145	2	F82133	hypothetical prote	473	31	42.5	385	2	F69941	conserved hypotet
401	31	42.5	146	2	T29401	hypothetical prote	474	31	42.5	389	2	C82637	porphyrin biosynth
402	31	42.5	147	2	T18586	hypothetical prote	475	31	42.5	405	2	T27509	homeotic protein 1
403	31	42.5	157	2	B29655	gamma-crystallin I	476	31	42.5	411	2	G69185	collagenase - Meth
404	31	42.5	167	2	T35368	hypothetical prote	477	31	42.5	411	2	T13315	hypothetical prote
405	31	42.5	170	2	AD0114	conserved hypotet	478	31	42.5	417	2	T33376	hypothetical prote
406	31	42.5	170	2	AG2479	aminoglycoside N6'	479	31	42.5	418	2	G84164	hypothetical prote
407	31	42.5	172	2	D72368	hypothetical prote	480	31	42.5	428	2	D97010	5-enolpyruvylshiki
408	31	42.5	173	2	S04266	gamma-crystallin I	481	31	42.5	428	2	A89759	seryl-tRNA synthet
409	31	42.5	173	2	C82062	hypothetical prote	482	31	42.5	429	2	S48172	variable surface g
410	31	42.5	182	2	JC7515	VPS29-like phospho	483	31	42.5	434	2	T47575	hypothetical prote
411	31	42.5	182	2	A52468	T-cell surface gly	484	31	42.5	454	2	A86221	hypothetical prote
412	31	42.5	188	2	H95211	anthranilate synth	485	31	42.5	459	2	A35291	adenylosuccinate 1
413	31	42.5	188	2	B98076	anthranilate synth	486	31	42.5	461	2	T33747	hypothetical prote
414	31	42.5	190	2	JC4514	TATA-binding prote	487	31	42.5	462	2	A84327	photolyase/cryptoc
415	31	42.5	191	2	E75072	transcription init	488	31	42.5	467	2	A47388	serine/threonine p
416	31	42.5	191	2	D71093	probable TARA-bind	489	31	42.5	474	2	T29336	hypothetical prote
417	31	42.5	192	2	S42381	hypothetical prote	490	31	42.5	475	2	C96503	protein F9C16.7 [l
418	31	42.5	207	2	A75475	probable acetyltra	491	31	42.5	476	2	G90512	p60-like (mycoplas
419	31	42.5	222	2	S26994	cellulose 1,4-beta	492	31	42.5	483	2	A39676	protein kinase Sry
420	31	42.5	224	2	S28737	hypothetical prote	493	31	42.5	488	2	D82991	conserved hypotet
421	31	42.5	228	2	S26995	cellulose 1,4-beta	494	31	42.5	489	2	AF2800	ABC transporter su
422	31	42.5	228	2	T47425	NAC domain-like pr	495	31	42.5	497	2	B97555	hypothetical prote
423	31	42.5	236	2	S26993	cellulose 1,4-beta	496	31	42.5	497	2	AD2775	fumarate hydratase
424	31	42.5	240	2	E86418	probable eukaryoti	497	31	42.5	501	2	T37747	hypothetical prote
425	31	42.5	240	2	T23797	hypothetical prote	498	31	42.5	503	2	AC1042	probable DNA helic
426	31	42.5	245	2	G82175	triose-phosphate i	499	31	42.5	510	2	S41943	cellulose 1,4-beta
427	31	42.5	250	1	ISUTTB	probable chaperrone	500	31	42.5	511	2	S44716	cellulose 1,4-beta
428	31	42.5	252	2	D91031	probable fibrial							
429	31	42.5	252	2	E85875	triose-phosphate i							
430	31	42.5	253	1	ISZMT	hypothetical prote							
431	31	42.5	253	2	A32187	triose-phosphate i							
432	31	42.5	260	2	B70419	hypothetical prote							
433	31	42.5	263	2	H82134	conserved hypotet							
434	31	42.5	267	2	JN0845	enterohemolysin 1							
435	31	42.5	268	2	T31699	phosphomethylpyrim							
436	31	42.5	271	2	AF1157	phosphomethylpyrim							
437	31	42.5	271	2	AC1516	protein kinase (EC							
438	31	42.5	274	2	T10270	conjugal transfer							
439	31	42.5	284	2	AG3231	probable myb-relat							
440	31	42.5	285	2	G85016	neurotrophin-6 - s							
441	31	42.5	286	2	S50855	hypothetical prote							
442	31	42.5	287	2	T25064	dihydrodipicolinat							
443	31	42.5	289	2	E64330	hypothetical prote							
444	31	42.5	290	2	AH2001	hypothetical prote							
445	31	42.5	292	2	T10106	chitinase (EC 3.2.							
446	31	42.5	293	2	AE3253	chromosome partiti							
447	31	42.5	306	2	I49068	protein kinase Sry							
448	31	42.5	307	2	T21879	hypothetical prote							
449	31	42.5	311	2	E96714	probable DNA-bind							
450	31	42.5	314	2	C90256	conserved hypotet							
451	31	42.5	317	2	T18695	hypothetical prote							
452	31	42.5	318	2	C93297	conserved hypotet							
453	31	42.5	319	2	T27700	hypothetical prote							
454	31	42.5	320	2	A86067	porphobilinogen de							
455	31	42.5	320	2	G91220	porphobilinogen de							
456	31	42.5	320	2	B44059	36K protein - appl							
457	31	42.5	321	2	S57482	chitinase class 1							
458	31	42.5	321	2	H87683	transcription regu							
459	31	42.5	323	2	T27640	hypothetical prote							
460	31	42.5	336	2	E69845	thiamin biosynthes							
461	31	42.5	338	2	H83228	hypothetical prote							
462	31	42.5	341	2	B95403	probable ABC trans							
463	31	42.5	342	2	A89846	hypothetical prote							
464	31	42.5	344	2	T27664	hypothetical prote							
465	31	42.5	347	2	JC7828	glucuronyltransfer							
466	31	42.5	359	2	AF3184	beta-lactamase [im							
467	31	42.5	364	2	E69131	tryptophan-tRNA li							

ALIGNMENTS

RESULT 1

C64398

hypothetical protein homolog MJ0787 - Methanococcus jannaschii

C:Species: Methanococcus jannaschii

C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004

C:Accession: C64398

R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, C.; Roon, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.

A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.

A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii.

A:Reference number: A64300; MUID:96337999; PMID:8688087

A:Accession: C64398

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-504 <BUL>

A:Cross-references: UNIPROT:Q58197; GB:U67523; GB:L77117; NID:g2826319; PIDN:AAB98783.1;

C:Genetics:

A:Map position: FOR710775-712289

Query Match 60.3%; Score 44; DB 2; Length 504;

Best Local Similarity 66.7%; Pred. No. 5.4;

Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 YRCRGDDSKVQE 13

DB 463 YPCNGDDKKVLE 474

RESULT 2

A70669

probable acid-CoA ligase (EC 6.2.1.-) fadD22 [similarity] - Mycobacterium tuberculosis (H37Rv)

N:Alternate names: hypothetical protein pks002a

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
 C/Accession: A70669; S73073
 R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
 A/Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A/Reference number: A70669
 A/Accession: A70669
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-705 <COL>
 A/Cross-references: UNIPROT:P96283; GB:Z83858; GB:AL123456; NID:g3261675; PIDN:CAB06101.
 R/Smith, D.R.; Robison, K.
 A/Description: Mycobacterium tuberculosis cosmid tbc2.
 A/Reference number: S73053
 A/Accession: S73073
 A/Molecule type: DNA
 A/Residues: 1-486, 'QQ', 489-705 <SMI>
 A/Cross-references: EMBL:U00024; NID:g560506; PIDN:AAA50930.1; PID:g560509
 C/Genetics:
 A/Gene: fadD22
 C/Keywords: acid-thiol ligase; carrier protein; phosphopantetheine; phosphoprotein
 F:51-486/Domain: acetate-CoA ligase homology <ACL>
 F:544-615/Domain: acyl carrier protein homology <ACPI>
 F:579/Binding site: phosphopantetheine (Ser) (covalent) #status predicted
 Query Match 60.3%; Score 44; DB 2; Length 705;
 Best Local Similarity 70.0%; Pred. No. 7.3;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 2 YRCRGDSDSKV 11
 |||||
 Db 383 YRCRADDTTEV 392
 |||||
 RESULT 3
 B87484
 anthranilate synthase component I [imported] - Caulobacter crescentus
 C/Species: Caulobacter crescentus
 C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
 C/Accession: B87484
 R/Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J. B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A/Title: Complete Genome Sequence of Caulobacter crescentus.
 A/Reference number: B87249; MUID:21173698; PMID:11259647
 A/Accession: B87484
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-513 <STO>
 A/Cross-references: UNIPROT:Q9A731; GB:AE005673; NID:gl3423344; PIDN:AAK23870.1; GSPDB:G
 C/Genetics:
 C/Supfamily: anthranilate synthase component I
 Query Match 57.5%; Score 42; DB 2; Length 513;
 Best Local Similarity 50.0%; Pred. No. 12;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 QY 2 YRCRGDSDSKVQE 13
 :|||||
 Db 73 WRCRGDAETAE 84
 |||||
 RESULT 4
 G83840
 ferric uptake regulation protein BH1527 [imported] - Bacillus halodurans (strain C-125)
 C/Species: Bacillus halodurans

C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
 C/Accession: G83840
 R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirano, T.; Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirano, T. Nucleic Acids Res. 28, 4317-4331, 2000
 A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
 A/Reference number: A83650; MUID:20512582; PMID:11058132
 A/Accession: G83840
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-150 <STO>
 A/Cross-references: UNIPROT:Q9KCP2; GB:AP001512; GB:BA000004; NID:gl0174030; PIDN:BA00524
 A/Experimental source: strain C-125
 C/Genetics:
 A/Gene: BH1527
 C/Supfamily: ferric uptake regulator
 Query Match 54.8%; Score 40; DB 2; Length 150;
 Best Local Similarity 66.7%; Pred. No. 9;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 2 YRCRGDSDSK 10
 :|||
 Db 141 HRCQGDSEK 149
 |||||
 RESULT 5
 T22521
 hypothetical protein F52H3.5 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C/Accession: T22521
 R/Gardner, A.
 A/Title: submitted to the EMBL Data Library, October 1995
 A/Reference number: Z19575
 A/Accession: T22521
 A/Status: preliminary; translated from GB/EMBL/DBBJ
 A/Molecule type: DNA
 A/Residues: 1-179 <WIL>
 A/Cross-references: UNIPROT:Q20683; EMBL:Z66512; PIDN:CAA91325.1; GSPDB:GN00020; CESP:F52
 A/Experimental source: clone F52H3
 C/Genetics:
 A/Gene: CESP:F52H3.5
 A/Map position: 2
 A/Introns: 88/2; 95/1; 135/3
 Query Match 54.8%; Score 40; DB 2; Length 179;
 Best Local Similarity 63.6%; Pred. No. 11;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 2 YRCRGDSDSKVQ 12
 |||||
 Db 125 YRLRGDDDKAR 135
 |||||
 RESULT 6
 T29321
 hypothetical protein M01E11.4 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
 C/Accession: T29321
 R/Pauley, A.; Gattung, S.
 A/Title: submitted to the EMBL Data Library, November 1996
 A/Description: The sequence of C. elegans cosmid M01E11.
 A/Reference number: Z20605
 A/Accession: T29321
 A/Status: preliminary; translated from GB/EMBL/DBBJ
 A/Molecule type: DNA
 A/Residues: 1-303 <PAU>
 A/Cross-references: EMBL:U00450; PIDN:AAB37827.1; GSPDB:GN00019; CESP:M01E11.4
 A/Experimental source: strain Bristol N2; clone M01E11
 C/Genetics:
 A/Gene: CESP:M01E11.4
 A/Map position: 1

A:introns: 32/3; 135/2; 152/2; 267/3
 C:Superfamily: Caenorhabditis elegans hypothetical protein M01E11.4

Query Match 54.8%; Score 40; DB 2; Length 303;
 Best Local Similarity 70.0%; Pred. No. 17;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 CRGDDSKVOE 13
 ||| ||| |||
 Db 42 CRGLDSQIOE 51

RESULT 7

A39130
 nicotinate phosphoribosyltransferase (EC 2.4.2.11) - Salmonella typhimurium

C:Species: Salmonella typhimurium
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C:Accession: A39130

R:Vinitsky, A.; Teng, H.; Grubmeyer, C.T.

J. Bacteriol. 173, 536-540, 1991

A:Title: Cloning and nucleic acid sequence of the Salmonella typhimurium pncB gene and a
 A:Reference number: A39130; MUID:91100340; PMID:1987148

A:Accession: A39130

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-400 <VIN>

A:Cross-references: UNIPROT:P22253; GB:M55986; NID:gl54268; PIDN:AAA27190.1; PID:gl54269

C:Genetics:

A:Gene: pncB

C:Function:

A:Pathway: nicotinate and nicotinamide metabolism

C:Superfamily: nicotinate phosphoribosyltransferase

C:Keywords: glycosyltransferase; pentosyltransferase

Query Match 54.8%; Score 40; DB 1; Length 400;
 Best Local Similarity 85.7%; Pred. No. 22;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDD 8
 : ||| |||
 Db 39 FRCRGDD 45

RESULT 8

JQ0756
 nicotinate phosphoribosyltransferase (EC 2.4.2.11) - Escherichia coli (strain K-12)

C:Species: Escherichia coli

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C:Accession: JQ0756; B64833

R:Wubbolter, M.G.; Terpstra, P.; van Beilen, J.B.; Kingma, J.; Meesters, H.A.R.; Witholt,
 J. Biol. Chem. 265, 17665-17672, 1990

A:Title: Variation of cofactor levels in Escherichia coli; sequence analysis and expres
 A:Reference number: JQ0756; MUID:91009224; PMID:2211655

A:Accession: JQ0756

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-400 <WUB>

A:Cross-references: UNIPROT:P18133; GB:J05568; NID:gl47306; PIDN:AAA24400.1; PID:gl47307

A:Experimental source: strain GEC70

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
 A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: B64833

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-400 <BLAT>

A:Cross-references: GB:AE000195; GB:U00096; NID:gl787156; PIDN:AACT4017.1; PID:gl787162;
 A:Experimental source: strain K-12, substrain MG1655

C:Genetics:

A:Gene: pncB

C:Function:

A:Pathway: nicotinate and nicotinamide metabolism
 C:Superfamily: nicotinate phosphoribosyltransferase
 C:Keywords: glycosyltransferase; pentosyltransferase

Query Match 54.8%; Score 40; DB 1; Length 400;
 Best Local Similarity 85.7%; Pred. No. 22;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDD 8
 : ||| |||
 Db 39 FRCRGDD 45

RESULT 9

AB0617
 nicotinate phosphoribosyltransferase [imported] - Salmonella enterica subsp. enterica ser

C:Species: Salmonella enterica subsp. enterica serovar Typhi

A:Note: this species has also been called Salmonella typhi

C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002

C:Accession: AB0617

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
 th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
 S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
 A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serova

A:Reference number: AB0502; MUID:21534947; PMID:11677608

A:Accession: AB0617

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-400 <PAR>

A:Cross-references: GB:AL513382; PIDN:CAD05404.1; PID:gl6502165; GSPDB:GN00176

C:Genetics:

A:Gene: STY1010

C:Superfamily: nicotinate phosphoribosyltransferase

Query Match 54.8%; Score 40; DB 2; Length 400;
 Best Local Similarity 85.7%; Pred. No. 22;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDD 8
 : ||| |||
 Db 39 FRCRGDD 45

RESULT 10

F90755
 nicotinate phosphoribosyltransferase [imported] - Escherichia coli (strain O157:H7, subst

C:Species: Escherichia coli

C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004

C:Accession: F90755

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
 Gasaawa, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hatcori, M.; Shinsagawa, H.

gasaawa, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hatcori, M.; Shinsagawa, H.

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genon
 A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: F90755

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-400 <HAY>

A:Cross-references: UNIPROT:Q8XDE8; GB:BA000007; PIDN:BA834437.1; PID:gl3360473; GSPDB:GN
 A:Experimental source: strain O157:H7, substrain RIMD 0509952

C:Genetics:

A:Gene: ECs1014

C:Superfamily: nicotinate phosphoribosyltransferase

Query Match 54.8%; Score 40; DB 2; Length 400;
 Best Local Similarity 85.7%; Pred. No. 22;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDD 8
 : ||| |||
 Db 39 FRCRGDD 45

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-721 <STM>
A;Cross-references: UNIPROT:Q9PA43; GB:AE004073; GB:AE003849; NID:g9107904; PIDN:AAF85474;
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Al-Jarrah, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H.; as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohme, J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigret, C.; Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.P.; Marino, C.L.; Marques, M.V.; Martins, E.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Miyaki, C.Y.; Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasaki, A.; Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira, M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Zé
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF2677

Query Match 54.8%; Score 40; DB 2; Length 721;
Best Local Similarity 54.5%; Pred. No. 39;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 YRCRGDSKVQ 12
| ||||| : :
Db 282 YACRGNDGKLK 292

RESULT 14
T43011
suppressor protein homolog - fission yeast (Schizosaccharomyces pombe) (fragment)
C;Species: Schizosaccharomyces pombe
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C;Accession: T43011
R;Yoshioka, S.; Kato, K.; Nakai, K.; Okayama, H.; Nojima, H.
DNA Res. 4, 363-369, 1997
A;Title: Identification of open reading frames in Schizosaccharomyces pombe cDNAs.
A;Reference number: Z17323; MUID:98162722; PMID:9501991
A;Accession: T43011
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-280 <YOS>
A;Cross-references: UNIPROT:P78857; EMBL:D89207; NID:gi749621; PIDN:BAAL3868.1; PID:gl714;
A;Experimental source: strain PR745
C;Superfamily: suppressor 2 protein; translation elongation factor Tu homology

Query Match 53.4%; Score 39; DB 2; Length 280;
Best Local Similarity 80.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 RCRGDSKVQ 12
| ||||| : :
Db 151 RVRGDSKVQ 160

RESULT 15
S66668
hydrogen peroxide-inducible protein - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C;Accession: S66668
R;Egawa, K.; Yoshiwara, M.; Shibamura, M.; Nose, K.
FEBS Lett. 372, 74-77, 1995
A;Title: Isolation of a novel ras-recision gene that is induced by hydrogen peroxide from
A;Reference number: S66668; MUID:96032549; PMID:7556647
A;Accession: S66668
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-325 <EGA>
A;Cross-references: UNIPROT:O61469; EMBL:L43371; NID:q1161099; PIDN:AAA85353.1; PID:ql161

A,Note: in the authors' translation residues 241-260 do not match the nucleotide sequend

Query Match 53.4%; Score 39; DB 2; Length 325;
Best Local Similarity 46.2%; Pred. No. 28;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYRCRGDDSKVQ 13
: || :||: ||:|
Db 261 DYICQNEKVE 273

RESULT 16
AD3382
ABC transporter-associated protein BMEI1042 [imported] - Brucella melitensis (strain 16M
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C:Accession: AD3382
R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltzman, E.; Selkov, E.; Elizer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:11756688
A:Accession: AD3382
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-507 <KUR>
A:Cross-references: UNIPROT:Q8YGW3; GB:AE008917; PIDN:AAL52223.1; PID:gi7983006; GSPDB:C
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI1042
A:Map position: I
A:Superfamily: conserved hypothetical protein b1683

Query Match 53.4%; Score 39; DB 2; Length 507;
Best Local Similarity 75.0%; Pred. No. 42;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 CRGDDSKV 11
|||:|:
Db 319 CRGDSKI 326

RESULT 17
T41442
omnipotent nonsense suppressor, efl alpha factor-like gtp-bindingprotein - fission yeast
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
R:Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, October 1998
A:Reference number: Z21993
A:Accession: T41442
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-662 <SEE>
A:Cross-references: UNIPROT:O74718; EMBL:AL032824; PIDN:CAA21821.1; GSPDB:GN00068; SPDB:
A:Experimental source: strain 972h-; cosmid c584
C:Genetics:
A:Gene: SPDB:SPCC584.04
A:Map position: 3
A:Introns: 44/1
A:Superfamily: suppressor 2 protein; translation elongation factor Tu homology
F:239-387/Domain: translation elongation factor Tu homology <ETU>

Query Match 53.4%; Score 39; DB 2; Length 662;
Best Local Similarity 80.0%; Pred. No. 54;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 RCRGDDSKVQ 12
| ||||| ||
Db 533 RVRGDDSDVQ 542

RESULT 18
T51948
omnipotent nonsense suppressor SUP35/erF-3 - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 09-Jul-2004
C:Accession: T51948
R:Ito, K.; Ebihara, K.; Nakamura, Y.
RNA 4, 958-972, 1998
A:Title: The stretch of C-terminal acidic amino acids of translational release factor erF
A:Reference number: Z25883
A:Accession: T51948
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-662 <ITO>
A:Cross-references: UNIPROT:O74718; EMBL:D79214; PIDN:BAA33530.1
A:Experimental source: strain JY333
C:Genetics:
A:Gene: sup35
A:Introns: 44/1
C:Superfamily: suppressor 2 protein; translation elongation factor Tu homology

Query Match 53.4%; Score 39; DB 2; Length 662;
Best Local Similarity 80.0%; Pred. No. 54;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 RCRGDDSKVQ 12
| ||||| ||
Db 533 RVRGDDSDVQ 542

RESULT 19
T14593
syringomycin synthetase - Pseudomonas syringae pv. syringae
C:Species: Pseudomonas syringae pv. syringae
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T14593
R:Guenzi, E.; Galli, G.; Grgurina, I.; Gross, D.C.; Grandi, G.
J. Biol. Chem. 273, 32857-32863, 1998
A:Title: Characterization of the syringomycin synthetase gene cluster. A link between prc
A:Reference number: Z18153; MUID:99047670; PMID:9830033
A:Accession: T14593
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-9376 <GUE>
A:Cross-references: UNIPROT:O85168; EMBL:AF047828; NID:g3510628; PID:g3510629; PIDN:AAC8(
C:Genetics:
A:Gene: syzE

C:Keywords: carrier protein; phosphopantetheine; phosphoprotein
F:441-896/Domain: acetate-CoA ligase homology <ACL1>
F:914-981/Domain: acyl carrier protein homology <ACP1>
F:1529-1984/Domain: acetate-CoA ligase homology <ACL2>
F:2002-2069/Domain: acyl carrier protein homology <ACP2>
F:2613-3071/Domain: acetate-CoA ligase homology <ACL3>
F:3089-3156/Domain: acyl carrier protein homology <ACP3>
F:3700-4158/Domain: acetate-CoA ligase homology <ACL4>
F:4176-4244/Domain: acyl carrier protein homology <ACP4>
F:4768-5224/Domain: acetate-CoA ligase homology <ACL5>
F:5242-5310/Domain: acyl carrier protein homology <ACP5>
F:5834-6280/Domain: acetate-CoA ligase homology <ACL6>
F:6298-6366/Domain: acyl carrier protein homology <ACP6>
F:6892-7352/Domain: acetate-CoA ligase homology <ACL7>
F:7370-7437/Domain: acyl carrier protein homology <ACP7>
F:7950-8440/Domain: acetate-CoA ligase homology <ACL8>
F:8458-8526/Domain: acyl carrier protein homology <ACP8>
F:9015-9083/Domain: acyl carrier protein homology <ACP>
F:946,2034,3121,4208,5274,6330/Binding site: phosphopantetheine (Ser) (covalent) #status

Query Match 53.4%; Score 39; DB 2; Length 9376;
Best Local Similarity 50.0%; Pred. No. 6.5e+02;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDDSKVQ 12
||| :||:|:

```

Db      3548 NYRHQGEDNRLQ 3559

RESULT 20
T31420
C:Terminal domain-binding protein rA8 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T31420
R:Yuryev, A.; Patturajan, M.; Litington, Y.; Joshi, R.V.; Gentile, C.; Gebara, M.; Cord
Proc. Natl. Acad. Sci. U.S.A. 93, 6975-6980, 1996
A:Title: The C-terminal domain of the largest subunit of RNA polymerase II interacts wit
A:Reference number: Z21024; MUID:96293459; PMID:8692929
A:Accession: T31420
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1268 <YUR>
A:Cross-references: UNIPROT:Q63623; EMBL:U49055; NID:g1438530; PIDN:AA5C26
A:Experimental source: hippocampus

Query Match      52.7%; Score 38.5; DB 2; Length 1268;
Best Local Similarity 72.7%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY      2 YRCGDD-SKV 11
      ||| ||| |||
Db      100 YRCFGDKSKI 110

RESULT 21
F81439
translation initiation factor IF-3 Cj0207 [imported] - Campylobacter jejuni (strain NCTC
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: F81439
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillin
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrel
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
A:Reference number: A81250; MUID:20150912; PMID:10688204
A:Accession: F81439
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-172 <PAR>
A:Cross-references: UNIPROT:Q9PI22; GB:AL139074; GB:AL111168; NID:g6967505; PIDN:CAB7269
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: InfC; Cj0207
C:Superfamily: translation initiation factor IF-3

Query Match      52.1%; Score 38; DB 2; Length 172;
Best Local Similarity 77.8%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      3 RCRGDDSKV 11
      ||| ||| |||
Db      19 RCVGDDGKV 27

RESULT 22
T27085
hypothetical protein Y51A2D.17 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 16-Aug-2004
C:Accession: T27085; T27320
R:McMurray, A.
submitted to the EMBL Data Library, January 1998
A:Reference number: Z20307
A:Accession: T27085
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-389 <WIL>
A:Cross-references: UNIPROT:Q9XTG0; UNIPROT:Q9GTD6; EMBL:AL021497; PIDN:CAA16413.1; GSPD

```

```

A:Experimental source: clone Y51A2D
R:McMurray, A.
submitted to the EMBL Data Library, August 1997
A:Reference number: Z20343
A:Accession: T27320
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-389 <W12>
A:Cross-references: EMBL:Z98877; PIDN:CAB11571.1; GSPDB:GN00023; CESP:Y51A2D.17
A:Experimental source: clone Y69H2
C:Genetics:
A:Gene: CESP:Y51A2D.17
A:Map position: 5
A:Superfamily: erBA transforming protein homology

Query Match      52.1%; Score 38; DB 2; Length 389;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 YRCRGD 7
      ||| |||
Db      45 YRCRGD 50

RESULT 23
E85343
hypothetical protein AT4G29440 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: E85343
R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: E85343
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1071 <STO>
A:Cross-references: UNIPROT:Q9MOD9; GB:NC_001268; NID:g7269843; PIDN:CAB79702.1; GSPDB:GN
C:Genetics:
A:Gene: AT4G29440
A:Map position: 4

Query Match      52.1%; Score 38; DB 2; Length 1071;
Best Local Similarity 77.8%; Pred. No. 1.3e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      5 RGDSSKVQE 13
      : ||| ||| |||
Db      640 QGDDSKTQE 648

RESULT 24
MWXR31
lambda 3 protein - reovirus type 1 (strain Lang)
N:Alternate names: minor core protein
C:Species: reovirus type 1
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
C:Accession: A30121
R:Wiener, J.R.; Joklik, W.K.
Virology 169, 194-203, 1989
A:Title: The sequences of the reovirus serotype 1, 2, and 3 L1 genome segments and analy
A:Reference number: A94390; MUID:89163254; PMID:2922925
A:Accession: A30121
A:Molecule type: genomic RNA
A:Residues: 1-1267 <WIE>
A:Cross-references: UNIPROT:P17376; GB:M24734; NID:g499863
A>Note: this sequence, which matches the sequence attributed to type 1 in Fig. 2, matched
he translations in entries REO1LAM3P and REO3LAM3P now differ only by the sequence corre
C:Comment: See also PIR:MWXR33.
A:Map position: segment L1

```

C:Superfamily: reovirus lambda 3 protein
C:Keywords: core protein

Query Match 52.1%; Score 38; DB 1; Length 1267;
Best Local Similarity 75.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYVRCGDD 8
||| |||
Db 728 NYVCGDD 735

RESULT 25

MWXR32

lambda 3 protein - reovirus type 2 (strain D5/Jones)

N:Alternate names: minor core protein

C:Species: reovirus type 2

A:Note: host Homo sapiens (man)

C>Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004

C:Accession: B30121

R:Wiener, J.R.; Joklik, W.K.

Virology 169, 194-203, 1989

A:Title: The sequences of the reovirus serotype 1, 2, and 3 L1 genome segments and analysis

A:Reference number: A94390; MUID:89163254; PMID:2922925

A:Accession: B30121

A:Molecule type: genomic RNA

A:Residues: 1-1267 <WIE>

A:Cross-references: UNIPROT:P17377; GB:M31057; NID:G499865; PIDN:AAA47245.1; PID:G499865

C:Genetics:

A:Map position: segment L1

C:Superfamily: reovirus lambda 3 protein

C:Keywords: core protein

Query Match 52.1%; Score 38; DB 1; Length 1267;
Best Local Similarity 75.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYVRCGDD 8
||| |||
Db 728 NYVCGDD 735

RESULT 26

MWXR33

lambda 3 protein - reovirus type 3 (strain Dearing)

N:Alternate names: minor core protein

C:Species: reovirus type 3

C>Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004

C:Accession: C30121

R:Wiener, J.R.; Joklik, W.K.

Virology 169, 194-203, 1989

A:Title: The sequences of the reovirus serotype 1, 2, and 3 L1 genome segments and analysis

A:Reference number: A94390; MUID:89163254; PMID:2922925

A:Accession: C30121

A:Molecule type: genomic RNA

A:Residues: 1-1267 <WIE>

A:Cross-references: UNIPROT:P17378; GB:M31058; NID:G499867; GB:M24734; NID:G499863; PIDN

A:Note: this sequence, which matches the sequence attributed to type 3 in Fig. 2, matches

varies R01LAM3P and R03LAM3P now differ only by the sequence correction apparently made

C:Comment: See also PIR:MWXR31.

C:Genetics:

A:Map position: segment L1

C:Superfamily: reovirus lambda 3 protein

C:Keywords: core protein

Query Match 52.1%; Score 38; DB 1; Length 1267;
Best Local Similarity 75.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYVRCGDD 8
||| |||
Db 728 NYVCGDD 735

RESULT 27

T28157

erythrocyte membrane protein 1 - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T28157

R:Yang, S.Q.; Wollish, W.S.; Gut, J.; Wu, J.; Ahn, J.; Petersen, C.; Fujioka, H.; Aikawa,

submitted to the EMBL Data Library, July 1995

A:Description: The molecular cloning and DNA sequence analysis of Plasmodium falciparum

A:Reference number: Z20479

A:Accession: T28157

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-2212 <YAN>

A:Cross-references: UNIPROT:Q94657; EMBL:U31083; NID:G1517813; PID:G1517814; PIDN:AAB0696

C:Genetics:

A:Gene: EMP1

Query Match 52.1%; Score 38; DB 2; Length 2212;
Best Local Similarity 45.5%; Pred. No. 2.5e+02;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 YRCRGDDSKVQ 12
::| ||| ::
Db 1770 FKCGGDDCRVR 1780

RESULT 28

T28161

hypothetical protein FCR3-varT11-1 - malaria parasite (Plasmodium falciparum) (fragments)

C:Species: Plasmodium falciparum

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T28161

R:Hernandez-Rivas, R.; Mattei, D.; Sterkers, Y.; Peterson, D.S.; Welles, T.E.; Scherf, F

Mol. Cell. Biol. 17, 604-611, 1997

A:Title: Expressed var genes are found in Plasmodium falciparum subtelomeric regions.

A:Reference number: Z20483; MUID:97154495; PMID:9001213

A:Accession: T28161

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-2647 <HER>

A:Cross-references: UNIPROT:P90580; EMBL:U67959; NID:G1794255; PID:G1809295; PIDN:AAC474;

A:Experimental source: strain FCQ27/PNG

C:Genetics:

A:Introns: 2158/3

A:Note: FCR3-varT11-1

Query Match 52.1%; Score 38; DB 2; Length 2647;
Best Local Similarity 45.5%; Pred. No. 3e+02;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 YRCRGDDSKVQ 12
::| ||| ::
Db 1778 FKCGGDDCRVR 1788

RESULT 29

T21381

hypothetical protein F26A3.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T21381

R:McMurray, A.

submitted to the EMBL Data Library, August 1996

A:Reference number: Z19415

A:Accession: T21381

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-3228 <WIL>

A:Cross-references: UNIPROT:Q93593; EMBL:Z78419; PIDN:CAB01701.1; GSPDB:GN00019; CESP:F26

A:Experimental source: clone F26A3

C:Genetics:

A:Gene: CESP:F26A3.3
A:Map position: 1

A:Introns: 36/3; 91/3; 216/3; 263/3; 308/3; 702/3; 828/3; 866/3; 1013/3; 1171/3; 1231/3; 2443/1; 2595/3; 2663/1; 2753/3; 2813/3; 2859/1; 3052/3; 3136/1; 3176/1

Query Match 52.1%; Score 38; DB 2; Length 3228;
Best Local Similarity 46.2%; Pred. No. 3.6e+02;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 NYRCRGDDSKVQ 13

||||| :||:

Db 2836 NYRCEDPDQEVK 2848

RESULT 30

H87107

ribosome recycling factor [imported] - Mycobacterium leprae

C:Species: Mycobacterium leprae

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C:Accession: H87107

R:Coile, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Holroyd, R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd, R.; M.A.; Rutherford, K.M.

Nature 409, 1007-1011, 2001

A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Seeger, K.

A:Title: Massive gene decay in the leprosy bacillus.

A:Reference number: A86909; MUID:21128732; PMID:11234002

A:Accession: H87107

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-185 <STO>

A:Cross-references: UNIPROT:O33046; GB:AL450380; NID:gl3093386; PIDN:CAC30541.1; GSPDB:G

C:Genetics: fr

A:Gene: fr

C:Superfamily: ribosome releasing factor

Query Match 50.7%; Score 37; DB 2; Length 185;

Best Local Similarity 55.6%; Pred. No. 38;

Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 3 RCRGDDSKV 11

||:|:|:

Db 118 KCKGEDAKV 126

RESULT 31

T21959

hypothetical protein F38C2.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T21959

R:Mortimore, B.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19494

A:Accession: T21959

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-258 <WIL>

A:Cross-references: UNIPROT:O45489; EMBL:Z82267; PIDN:CAB05189.1; GSPDB:GN00022; CESP:F3

A:Experimental source: clone F38C2

C:Genetics:

A:Gene: CESP:F38C2.2

A:Map position: 4

A:Introns: 56/2; 100/3; 184/2

Query Match

Best Local Similarity 50.7%; Score 37; DB 2; Length 258;

Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 NYRCRGDDSKVQ 12

||:|:|:

Db 103 NYKWRIEDQKVQ 114

RESULT 32

H86244

lysophospholipase homolog, 25331-24357 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 12-Jul-2004

C:Accession: H86244

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Jensen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, R.

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: H86244

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-324 <STO>

A:Cross-references: UNIPROT:O04083; GB:AE005172; NID:gl3131639; PIDN:AAB65474.1; GSPDB:GN

C:Genetics:

A:Map position: 1

C:Superfamily: tropinesterase

Query Match 50.7%; Score 37; DB 2; Length 324;

Best Local Similarity 58.3%; Pred. No. 64;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 NYRCRGDDSKVQ 12

||:|:|:

Db 312 NDRCGDKTKTQ 323

RESULT 33

T21960

hypothetical protein F38C2.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T21960

R:Mortimore, B.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19494

A:Accession: T21960

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-410 <WIL>

A:Cross-references: UNIPROT:O45490; EMBL:Z82267; PIDN:CAB05190.1; GSPDB:GN00022; CESP:F3

A:Experimental source: clone F38C2

C:Genetics:

A:Gene: CESP:F38C2.4

A:Map position: 4

A:Introns: 63/1; 79/3; 103/1; 138/1; 220/2; 279/3; 346/3; 367/2

Query Match

Best Local Similarity 50.7%; Score 37; DB 2; Length 410;

Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 NYRCRGDDSKVQ 12

||:|:|:

Db 3 NYKWRIEDQKVQ 14

RESULT 34

S75313

serine-tRNA ligase (EC 6.1.1.11) - Synecocystis sp. (strain PCC 6803)

N:Alternate names: protein slr1703; seryl-tRNA synthetase

C:Species: Synecocystis sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004

C:Accession: S75313

R.Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp.

A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S75313

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-430 <KAN>

A:Cross-references: UNIPROT:P73201; EMBL:D50904; GB:AB001339; NID:G1522225; PIDN:BAAL1722

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Genetics:

A:Gene: sers

A:Start codon: GTG

C:Function:

A:Description: charges tRNA(Ser) with serine

A:Pathway: protein biosynthesis

C:Superfamily: serine-tRNA ligase

C:Keywords: aminoacyl-tRNA synthetase; ATP; ligase; protein biosynthesis

Query Match 50.7%; Score 37; DB 2; Length 430;
Best Local Similarity 62.5%; Pred. No. 83;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYRCRGDD 8
:::|||||

Db 225 SFQCRGDD 232

RESULT 35

C87480

conserved hypothetical protein CC1864 [imported] - *Caulobacter crescentus*

C:Species: *Caulobacter crescentus*

C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C:Accession: C87480

R.Nierman, W.C.; Feldblym, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.; B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon, J.; Esmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of *Caulobacter crescentus*.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: C87480

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-489 <STO>

A:Cross-references: UNIPROT:Q9A762; GB:AE005673; NID:G13423307; PIDN:AAK23839.1; GSPDB:C

C:Genetics:

A:Gene: CC1864

C:Superfamily: conserved hypothetical protein b1683

Query Match 50.7%; Score 37; DB 2; Length 489;
Best Local Similarity 87.5%; Pred. No. 94;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CRGDRSKV 11
|||||

Db 301 CRGDRSKV 308

RESULT 36

T19133

hypothetical protein C09F9.5b - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T19133

R.Smye, R.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19078

A:Accession: T19133

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-518 <WIL>

A:Cross-references: EMBL:Z81465; PIDN:CAB03864.1; GSPDB:GN00020; CESP:C09F9.5b

A:Experimental source: clone C09F9

C:Genetics:

A:Gene: CESP:C09F9.5b

A:Map position: 2

A:Introns: 34/2; 69/3; 139/3; 222/3; 284/2; 335/2; 377/1

Query Match 50.7%; Score 37; DB 2; Length 518;
Best Local Similarity 58.3%; Pred. No. 99;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 YRCRGDDSKVQE 13
|||||

Db 55 YRCFPDESQVYE 66

RESULT 37

T19132

hypothetical protein C09F9.5a - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T19132

R.Smye, R.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19078

A:Accession: T19132

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-564 <WIL>

A:Cross-references: EMBL:Z81465; PIDN:CAB03863.1; GSPDB:GN00020; CESP:C09F9.5a

A:Experimental source: clone C09F9

C:Genetics:

A:Gene: CESP:C09F9.5a

A:Map position: 2

A:Introns: 34/2; 69/3; 139/3; 222/3; 284/2; 335/2; 377/1; 505/1

Query Match 50.7%; Score 37; DB 2; Length 564;
Best Local Similarity 58.3%; Pred. No. 1.1e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 YRCRGDDSKVQE 13
|||||

Db 55 YRCFPDESQVYE 66

RESULT 38

A89862

Na+/H+ antiporter subunit [imported] - *Staphylococcus aureus* (strain N315)

C:Species: *Staphylococcus aureus*

C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004

C:Accession: A89862

R.Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguchi, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K. Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.

A:Reference number: A89758; MUID:21311952; PMID:11418146

A:Accession: A89862

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-801 <KUR>

A:Cross-references: UNIPROT:P60674; UNIPROT:P60675; GB:BA000018; PID:G13700756; PIDN:BAB4

A:Experimental source: strain N315

C:Genetics:

A:Gene: mnhA

C:Superfamily: NADH dehydrogenase (ubiquinone) chain 5

Query Match 50.7%; Score 37; DB 2; Length 801;
Best Local Similarity 54.5%; Pred. No. 1.5e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYRCRGDDSKV 11
:|:|||||

Db 322 SYHYQGDSDSKI 332

RESULT 39

T19131
 hypothetical protein C09F9.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T19131
 R:Smyle, R.
 submitted to the EMBL Data Library, November 1996
 A:Reference number: Z19078
 A:Accession: T19131
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-873 <WIL>
 A:Cross-references: EMBL:Z81465; PIDN:CAB03862.1; GSPDB:GN00020; CESP:C09F9.3
 A:Experimental source: clone C09F9
 C:Genetics:
 A:Gene: CESP.C09F9.3
 A:Map position: 2
 A:Introns: 66/1; 168/2; 293/1; 343/2; 378/3; 448/3; 531/3; 593/2; 644/2; 686/1; 814/1

Query Match 50.7%; Score 37; DB 2; Length 873;
 Best Local Similarity 58.3%; Pred. No. 1.6e+02;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 YRCRGDDSKVQE 13
 ||| | | | |
 Db 364 YRCFPDESQVYE 375

RESULT 40

F86316
 protein T10022.13 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C:Accession: F86316
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lurob, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I. ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: F86316
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-988 <STO>
 A:Cross-references: UNIPROT:Q9LM32; GB:AE005172; NID:g8671767; PIDN:AAF78373.1; GSPDB:GN
 C:Genetics:
 A:Gene: T10022.13
 A:Map position: 1

Query Match 50.7%; Score 37; DB 2; Length 988;
 Best Local Similarity 54.5%; Pred. No. 1.8e+02;
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 NYRCRGDDSKV 11
 ||| | | | |
 Db 164 NYNCLGVDDKI 174

RESULT 41

F96763
 hypothetical protein F25P22.8 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C:Accession: F96763

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lurob, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I. ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: F96763
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1030 <STO>
 A:Cross-references: UNIPROT:Q9C9U5; GB:AE005173; NID:g6692730; PIDN:AAF24836.1; GSPDB:GN
 C:Genetics:
 A:Gene: F25P22.8
 A:Map position: 1

Query Match 50.7%; Score 37; DB 2; Length 1030;
 Best Local Similarity 54.5%; Pred. No. 1.9e+02;
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 NYRCRGDDSKV 11
 ||| | | | |
 Db 188 NYNCLGVDDKI 198

RESULT 42

C90460
 hypothetical protein SSO2828 [imported] - Sulfolobus solfataricus
 C:Species: Sulfolobus solfataricus
 C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
 C:Accession: C90460
 R:She, O.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-v Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P. arett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
 submitted to GenBank, April 2001
 A:Description: Sulfolobus solfataricus complete genome.
 A:Reference number: A99139
 A:Accession: C90460
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-138 <KUR>
 A:Cross-references: UNIPROT:Q97V09; GB:AE006641; NID:gl3816183; PIDN:AAK42938.1; GSPDB:GN
 C:Genetics:
 A:Gene: SSO2828

Query Match 49.3%; Score 36; DB 2; Length 138;
 Best Local Similarity 60.0%; Pred. No. 43;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 YRCRGDDSKV 11
 ||| | | | |
 Db 45 YRCBGDIGKL 54

RESULT 43

AC3270
 hypothetical exported protein BMEI0144 [imported] - Brucella melitensis (strain 16M)
 C:Species: Brucella melitensis
 C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
 C:Accession: AC3270
 R:DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, I. ; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letesc Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
 A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
 A:Reference number: AD3252; PMID:11756688
 A:Accession: AC3270
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-180 <KUR>

Db 187 NVRFKGGDGAEK 199

RESULT 46

C85720

probable oxidoreductase ydfG [imported] - Escherichia coli (strain O157:H7, substrain EDI

C/Species: Escherichia coli

C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004

C/Accession: C85720

R/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, K.;

Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dinalanta, E.; Potamousis, K.; Apodaca,

Nature 409, 529-533, 2001

A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A/Reference number: A85480; PMID:21074935; PMID:11206551

A/Accession: C85720

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-248 <STO>

A/Cross-references: UNIPROT:Q8X505; GB:AE005174; NID:gi2515117; PIDN:AAG56223.1; GSPDB:F07

A/Experimental source: strain O157:H7, substrain EDL933

C/Genetics:

A/Gene: ydfG

C/Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 49.3%; Score 36; DB 2; Length 248;

Best Local Similarity 46.2%; Pred. No. 75;

Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 NVRCRGDDSKVOE 13

Db 187 NVRFKGGDGAEK 199

RESULT 47

T20546

hypothetical protein F07C6.1 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C/Accession: T20546

R/Steward, C.

submitted to the EMBL Data Library, February 1996

A/Reference number: Z19290

A/Accession: T20546

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-329 <WTL>

A/Cross-references: UNIPROT:Q19157; EMBL:Z69659; PIDN:CAA93483.1; GSPDB:GN00022; CBSP:F07

A/Experimental source: Clone F07C6

C/Genetics:

A/Gene: CBSP:F07C6.1

A/Map position: 4

A/Introns: 14/2; 47/2; 85/1; 146/1; 226/3; 299/2

Query Match 49.3%; Score 36; DB 2; Length 329;

Best Local Similarity 50.0%; Pred. No. 98;

Matches 9; Conservative 0; Mismatches 3; Indels 6; Gaps 1;

QY 2 YRCRGDD-----SKVOE 13

Db 287 YRCRGCDKVLHYKDKVME 304

RESULT 48

T35013

probable membrane protein - Streptomyces coelicolor

C/Species: Streptomyces coelicolor

C/Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004

C/Accession: T35013

R/Seeger, S.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, June 1999

A/Reference number: Z21565

A/Accession: T35013

A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-347 <SEE>
A;Cross-references: UNIPROT:Q9XAN5; EMBL:AL079355; PIDN:CAB45569.1; GSPDB:GN00070; SCOE
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SC4C6.05c
C;Superfamily: Streptomyces coelicolor probable membrane protein SC4C6.05c

Query Match 49.3%; Score 36; DB 2; Length 347;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 YRCRGDDSK 10
|:|:|:|:|
DB 206 YKCPGDTSK 214

RESULT 49
H72494
hypothetical protein APE2599 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: H72494
R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: H72494
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-411 <XAW>
A;Cross-references: UNIPROT:Q9Y8N4; DBJ:AP000064; NID:G5105945; PIDN:BAA81616.1; PID:G5
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE2599
C;Superfamily: Aeropyrum pernix hypothetical protein APE2599

Query Match 49.3%; Score 36; DB 2; Length 411;
Best Local Similarity 54.5%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 YRCRGDDSKVQ 12
|:|:|:|:|
DB 225 YNCRDEDNVVQ 235

RESULT 50
A46206
voltage-gated sodium channel alpha subunit - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A46206
R;Gautron, S.; Dos Santos, G.; Pinto-Henrique, D.; Koulakoff, A.; Gros, F.; Berwald-Nett
Proc. Natl. Acad. Sci. U.S.A. 89, 7272-7276, 1992
A;Title: The glial voltage-gated sodium channel: cell- and tissue-specific mRNA expressi
A;Reference number: A46206; MUID:92357813; PMID:1379737
A;Accession: A46206
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-435 <GAU>
A;Cross-references: UNIPROT:Q01340; GB:M96578; NID:G204544; PIDN:AAA41303.1; PID:G204545
A;Experimental source: 1-day-old Wistar, astrocyte cultures
A;Note: sequence extracted from NCBI backbone (NCBIP:113349)
C;Superfamily: sodium channel protein

Query Match 49.3%; Score 36; DB 2; Length 435;
Best Local Similarity 46.2%; Pred. No. 1.3e+02;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 NYRCRGDDSKVQE 13
|:|:|:|:|

Db 392 SYRLRQSDKKIQD 404

Search completed: September 7, 2005, 20:04:01
Job time : 14.9767 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2005, 19:43:14 ; Search time 45.6512 Seconds
(without alignments)
145.824 Million cell updates/sec

Title: US-10-812-238b-23

Perfect score: 73

Sequence: 1 NYRCRGDDSKVQE 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database : Uniprot_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	73	100.0	311	1 LPP3 HUMAN	O1495 h lipid pho
2	69	94.5	312	1 LPP3_MOUSE	Q991y8 mus musculus
3	69	94.5	312	1 LPP3_RAT	P97544 rattus norv
4	69	94.5	312	2 Q61MX4	O61mx4 rattus norv
5	46	63.0	307	2 Q68F34	Q68f34 xenopus lae
6	46	63.0	776	2 Q73RY0	Q73ry0 mycobacteri
7	44	60.3	167	2 Q95L84	Q95l84 oryctolagus
8	44	60.3	504	1 Y787_METJA	Q58197 methanococc
9	44	60.3	705	2 Q50468	Q50468 mycobacteri
10	44	60.3	705	2 P96283	P96283 mycobacteri
11	44	60.3	705	2 Q7TXK7	Q7txk7 mycobacteri
12	43	58.9	307	2 Q6AX87	Q6ax87 xenopus lae
13	42	57.5	73	2 Q7QP19	Q7qp19 giardia lam
14	42	57.5	265	2 Q86KQ7	Q86kq7 dictyosteli
15	42	57.5	513	2 Q9A731	Q9a731 caulobacter
16	42	57.5	845	2 Q8WQJ3	Q8wgj3 plasmodium
17	42	57.5	1304	2 Q8WT63	Q8wt63 plasmodium
18	41	56.2	37	2 Q8POM2	Q8pom2 streptococ
19	41	56.2	226	2 Q69XV4	Q69xv4 oryza sativ
20	41	56.2	428	2 Q8TBF1	Q8tbf1 homo sapien
21	41	56.2	461	2 Q69XV5	Q69xv5 oryza sativ
22	41	56.2	489	2 Q8N9B8	Q8n9b8 homo sapien
23	41	56.2	505	2 Q96I80	Q96i80 homo sapien
24	41	56.2	875	1 NETE_HUMAN	P56730 homo sapien
25	41	56.2	1334	2 Q97322	Q97322 plasmodium
26	41	56.2	2736	2 Q7RTC3	Q7rtc3 plasmodium
27	40	54.8	150	2 Q9KCP2	Q9kcp2 bacillus ha
28	40	54.8	179	2 Q20683	Q20683 caenorhabdi
29	40	54.8	244	2 Q95ZL6	Q95z16 caenorhabdi
30	40	54.8	252	2 Q95ZL7	Q95z17 caenorhabdi
31	40	54.8	269	2 Q7R8M3	Q7r8m3 plasmodium

32	40	54.8	282	1 LPP1_RAT	O08564 rattus norv
33	40	54.8	282	2 Q6P766	Q6p766 rattus norv
34	40	54.8	288	2 Q9DW23	Q9dw23 rat cytomog
35	40	54.8	304	1 HEN3_XANCP	Q8p536 xanthomonas
36	40	54.8	309	1 PNCB_ECO57	Q8p536 escherichia
37	40	54.8	399	1 PNCB_ECOLI	P18133 escherichia
38	40	54.8	399	1 PNCB_SALTI	Q8t7y9 salmonella
39	40	54.8	399	1 PNCB_SALTY	P22253 salmonella
40	40	54.8	400	2 Q7UD27	Q7ud27 shigella fl
41	40	54.8	416	2 Q83LN3	Q83ln3 shigella fl
42	40	54.8	416	1 Q8FJ98	Q8fj98 escherichia
43	40	54.8	577	1 YQ4B_CAEEL	Q17426 caenorhabdi
44	40	54.8	586	2 Q62YB1	Q62yb1 bacillus li
45	40	54.8	587	2 Q65MW8	Q65mw8 bacillus li
46	40	54.8	721	2 Q9PA43	Q9pa43 xylella fas
47	40	54.8	1129	2 Q7RQD4	Q7rkd4 rhodospirell
48	40	54.8	1615	2 Q7RQM4	Q7rgm4 plasmodium
49	39	53.4	46	2 Q848W4	Q848w4 bacillus me
50	39	53.4	154	2 Q8BUS2	Q8bus2 mus musculu
51	39	53.4	278	2 Q6DHF1	Q6dhf1 brachydanio
52	39	53.4	280	2 P78857	P78857 schizosacch
53	39	53.4	283	1 LPP1_MOUSE	Q61469 mus musculu
54	39	53.4	305	2 Q9RDP5	Q9rdp5 streptomyce
55	39	53.4	403	2 Q6MQE4	Q6mqe4 bdellovibri
56	39	53.4	503	2 Q6FZW1	Q6fzw1 bartonella
57	39	53.4	503	2 Q6G3B9	Q6g3b9 bartonella
58	39	53.4	507	2 Q8YGW3	Q8ygw3 bruceella me
59	39	53.4	507	2 Q8G100	Q8g100 bruceella su
60	39	53.4	528	2 Q97379	Q97379 strongyloce
61	39	53.4	540	2 Q76771	Q76771 dictyosteli
62	39	53.4	588	2 Q70681	Q70681 sugarcane s
63	39	53.4	662	1 ERF2_SCHPO	Q74718 schizosacch
64	39	53.4	750	2 Q26188	Q26188 plasmodium
65	39	53.4	1015	2 Q6C6V1	Q6c6v1 yarrowia li
66	39	53.4	1420	2 Q9YNB1	Q9ynb1 sugarcane s
67	39	53.4	1450	2 Q815X9	Q815x9 plasmodium
68	39	53.4	1778	2 Q9BPP7	Q9bpb7 anthersea y
69	39	53.4	1778	2 Q9GUX5	Q9gux5 antheraea p
70	39	53.4	9376	2 Q85168	Q85168 pseudomonas
71	38.5	52.7	1268	2 Q63623	Q63623 rattus norv
72	38.5	52.7	1268	2 Q6D1D3	Q6did3 mus musculu
73	38.5	52.7	1271	1 RBMG_HUMAN	Q6n66 homo sapien
74	38.5	52.7	1271	2 Q6NSK3	Q6nsk3 homo sapien
75	38.5	52.7	1362	2 Q80TJ3	Q80tj3 mus musculu
76	38.5	52.7	1364	2 Q8I3E6	Q8i3e6 plasmodium
77	38	52.1	121	2 Q8Q0W8	Q8q0w8 methanosarc
78	38	52.1	121	2 Q8V163	Q8vi63 reovirus sp
79	38	52.1	121	2 Q8V164	Q8vi64 reovirus sp
80	38	52.1	121	2 Q8V165	Q8vi65 reovirus sp
81	38	52.1	121	2 Q8V166	Q8vi66 reovirus sp
82	38	52.1	121	2 Q8V167	Q8vi67 reovirus sp
83	38	52.1	121	2 Q8V168	Q8vi68 reovirus sp
84	38	52.1	121	2 Q8V169	Q8vi69 reovirus sp
85	38	52.1	121	2 Q8V170	Q8vi70 reovirus sp
86	38	52.1	121	2 Q8V171	Q8vi71 reovirus sp
87	38	52.1	121	2 Q8V172	Q8vi72 reovirus sp
88	38	52.1	121	2 Q8V173	Q8vi73 reovirus sp
89	38	52.1	121	2 Q8V174	Q8vi74 reovirus sp
90	38	52.1	121	2 Q8V175	Q8vi75 reovirus sp
91	38	52.1	121	2 Q8V176	Q8vi76 reovirus sp
92	38	52.1	121	2 Q8V177	Q8vi77 reovirus sp
93	38	52.1	121	2 Q8V178	Q8vi78 reovirus sp
94	38	52.1	121	2 Q8V179	Q8vi79 reovirus sp
95	38	52.1	121	2 Q8V180	Q8vi80 reovirus sp
96	38	52.1	121	2 Q8V181	Q8vi81 reovirus sp
97	38	52.1	121	2 Q8V182	Q8vi82 reovirus sp
98	38	52.1	121	2 Q8V183	Q8vi83 reovirus sp
99	38	52.1	121	2 Q8V184	Q8vi84 reovirus sp
100	38	52.1	121	2 Q8V185	Q8vi85 reovirus sp
101	38	52.1	121	2 Q8V186	Q8vi86 reovirus sp
102	38	52.1	121	2 Q8V187	Q8vi87 reovirus sp
103	38	52.1	121	2 Q8V188	Q8vi88 reovirus sp
104	38	52.1	121	2 Q8V189	Q8vi89 reovirus sp

105	38	52.1	121	2	Q8V190	Q8v190 reovirus sp	178	37	50.7	467	1	VE2_HPV24	P50770 human papil
106	38	52.1	121	2	Q8V191	Q8v191 reovirus sp	179	37	50.7	485	2	Q8P748	Q8p748 xanthomonas
107	38	52.1	121	2	Q8V192	Q8v192 reovirus sp	180	37	50.7	489	2	Q8A762	Q8a762 caulobacter
108	38	52.1	121	2	Q8V193	Q8v193 reovirus sp	181	37	50.7	490	2	Q82137	Q82137 streptomyce
109	38	52.1	121	2	Q8V194	Q8v194 reovirus sp	182	37	50.7	566	2	Q8S313	Q8s313 phytophthor
110	38	52.1	121	2	Q8V195	Q8v195 reovirus sp	183	37	50.7	743	2	Q8T1S1	Q8t1s1 methanosarc
111	38	52.1	121	2	Q8V196	Q8v196 reovirus sp	184	37	50.7	801	1	MNHA_STAAM	P60674 staphylococ
112	38	52.1	121	2	Q8V197	Q8v197 reovirus sp	185	37	50.7	801	1	MNHA_STAAM	P60675 staphylococ
113	38	52.1	121	2	Q8V198	Q8v198 reovirus sp	186	37	50.7	801	1	MNHA_STAAM	Q9nrg6 staphylococ
114	38	52.1	121	2	Q8V199	Q8v199 reovirus sp	187	37	50.7	801	1	MNHA_STAAM	Q8nxf6 staphylococ
115	38	52.1	121	2	Q8V200	Q8v200 reovirus sp	188	37	50.7	801	2	Q6GAX4	Q6gax4 staphylococ
116	38	52.1	121	2	Q8V201	Q8v201 reovirus sp	189	37	50.7	801	2	Q6GID6	Q6gid6 staphylococ
117	38	52.1	121	2	Q8V202	Q8v202 reovirus sp	190	37	50.7	808	2	Q7Q4C6	Q7q4c6 anopheles g
118	38	52.1	121	2	Q8V203	Q8v203 reovirus sp	191	37	50.7	850	2	Q6C6X5	Q6c6x5 yarrowia li
119	38	52.1	121	2	Q8V204	Q8v204 reovirus sp	192	37	50.7	868	2	Q6CFX2	Q6cfx2 yarrowia li
120	38	52.1	130	2	Q9YHG0	Q9yhg0 ginglymosto	193	37	50.7	888	2	Q75AE3	Q75ae3 ashbya goss
121	38	52.1	172	1	IF3_CAMJE	Q9pis2 campylobact	194	37	50.7	953	2	Q7PGN3	Q7pgn3 anopheles g
122	38	52.1	172	1	Q897T4	Q897t4 clostridium	195	37	50.7	966	2	Q9LMS1	Q9lms1 arabidopsis
123	38	52.1	191	2	Q9AY18	Q9ay18 funaria apo	196	37	50.7	970	2	Q7X8K7	Q7x8k7 brassica ju
124	38	52.1	193	2	Q9AY16	Q9ay16 entosthodon	197	37	50.7	988	2	Q9LM32	Q9lm32 arabidopsis
125	38	52.1	206	2	Q8TBK8	Q8tbk8 homo sapien	198	37	50.7	992	2	Q8LPH3	Q8lph3 arabidopsis
126	38	52.1	285	1	LP1_CAVPO	Q88956 cavia porce	199	37	50.7	1022	2	Q8XM34	Q8xm34 clostridium
127	38	52.1	298	2	Q7Q232	Q7q232 anopheles g	200	37	50.7	1030	2	Q8L625	Q8l625 arabidopsis
128	38	52.1	298	2	Q84KP4	Q84kp4 cyanidiosch	201	37	50.7	1030	2	Q9C9U5	Q9c9u5 arabidopsis
129	38	52.1	304	1	HEM3_XANAC	Q8dpr3 xanthomonas	202	37	50.7	1257	2	Q7QCZ7	Q7qc27 anopheles g
130	38	52.1	310	2	Q8D1T7	Q8dit7 synechococ	203	37	50.7	1325	2	Q95ZZ0	Q95zz0 caenorhabdi
131	38	52.1	311	2	Q8N4J2	Q8n4j2 homo sapien	204	37	50.7	1583	2	Q7RRC0	Q7rrc0 plasmodium
132	38	52.1	383	2	Q9GTD8	Q9gtd8 caenorhabdi	205	37	50.7	1756	2	Q8WPM5	Q8wpm5 oikopleura
133	38	52.1	387	2	Q9GTD7	Q9gtd7 caenorhabdi	206	37	50.7	1969	2	Q89HL6	Q89hl6 ciona inte
134	38	52.1	388	2	Q9GTD6	Q9gtd6 caenorhabdi	207	37	50.7	4569	2	Q7PS35	Q7ps35 anopheles g
135	38	52.1	389	2	Q9XTG0	Q9xtg0 caenorhabdi	208	36.5	50.0	459	2	Q7QJN6	Q7qjn6 anopheles g
136	38	52.1	435	2	Q8JTB2	Q8jtb2 golden ide	209	36.5	50.0	1193	2	Q7Z8L8	Q7z8l8 cochliobolu
137	38	52.1	454	2	Q9H648	Q9h648 homo sapien	210	36	49.3	63	2	Q9BS83	Q9bs83 homo sapien
138	38	52.1	476	2	Q6B8Q6	Q6b8q6 gracilaria	211	36	49.3	84	2	Q9RIJ0	Q9rij0 streptococ
139	38	52.1	542	2	Q8N3Z6	Q8n3z6 homo sapien	212	36	49.3	122	2	Q85KS8	Q85ks8 bacillus li
140	38	52.1	639	2	Q8DID1	Q8did1 synechococ	213	36	49.3	138	2	Q97V09	Q97v09 sulfolobus
141	38	52.1	661	2	Q8JTB6	Q8jtb6 striped bas	214	36	49.3	142	2	Q87I31	Q87i31 vibrio para
142	38	52.1	750	2	Q6CI30	Q6ci30 yarrowia li	215	36	49.3	147	1	Y237_TROWT	P67323 tropheryma
143	38	52.1	915	2	Q7UV66	Q7uv66 rhodospirell	216	36	49.3	147	1	Y533_TROW8	P67324 tropheryma
144	38	52.1	973	2	Q9NC89	Q9nc89 strongyloce	217	36	49.3	158	2	Q8LNX9	Q8lnx9 zinnia eleg
145	38	52.1	1071	2	Q9M0D9	Q9m0d9 arabidopsis	218	36	49.3	160	1	RRAL_VIBPA	Q87sd2 vibrio para
146	38	52.1	1075	2	Q9NC90	Q9nc90 strongyloce	219	36	49.3	164	2	Q7Q0D5	Q7q0d5 anopheles g
147	38	52.1	1240	2	Q8VA42	Q8va42 chum salmon	220	36	49.3	165	2	Q9GLE6	Q9gle6 sus scrofa
148	38	52.1	1267	1	VL3_REOVD	Pl7378 reovirus ty	221	36	49.3	176	2	Q8XN04	Q8xn04 clostridium
149	38	52.1	1267	1	VL3_REOVJ	Pl7377 reovirus ty	222	36	49.3	180	2	Q8VJE2	Q8vje2 bruceella me
150	38	52.1	1267	1	VL3_REOVL	Pl7376 reovirus ty	223	36	49.3	180	2	Q8FYG1	Q8fyg1 bruceella su
151	38	52.1	1267	2	Q8V5E5	Q8v5e5 ndelle viru	224	36	49.3	188	2	Q83XE8	Q83xe8 burkholderi
152	38	52.1	1267	2	Q85665	Q85665 reovirus ty	225	36	49.3	225	1	CLD8_HUMAN	P56748 homo sapien
153	38	52.1	1274	2	Q8JU61	Q8ju61 golden shin	226	36	49.3	225	1	CLD8_MOUSE	Q9z260 mus musculu
154	38	52.1	1274	2	Q9E3V9	Q9e3v9 grass carp	227	36	49.3	248	1	YDFG_ECOLI	P39831 escherichia
155	38	52.1	1601	2	Q9NDH1	Q9ndh1 caenorhabdi	228	36	49.3	248	2	Q62I15	Q62i15 burkholderi
156	38	52.1	212	2	Q9A657	Q9a657 plasmodium	229	36	49.3	248	2	Q63RB3	Q63rb3 burkholderi
157	38	52.1	2647	2	P90580	P90580 plasmodium	230	36	49.3	248	2	Q7UCH2	Q7uch2 shigella fl
158	37	50.7	59	2	Q6ETU2	Q6etu2 oryza sativ	231	36	49.3	248	2	Q8X505	Q8x505 escherichia
159	37	50.7	97	2	Q72N54	Q72n54 leptospira	232	36	49.3	250	2	Q6KX6	Q6kcx6 escherichia
160	37	50.7	97	2	Q8F8G5	Q8f8g5 leptospira	233	36	49.3	250	2	Q8CVU9	Q8cvu9 escherichia
161	37	50.7	108	2	Q6ESN2	Q6esn2 oryza sativ	234	36	49.3	252	2	Q937E7	Q937e7 nostoc punc
162	37	50.7	141	2	Q7MUA2	Q7muu2 porphyromon	235	36	49.3	253	2	Q83RE8	Q83re8 shigella fl
163	37	50.7	142	2	Q88Z98	Q88z98 lactobacill	236	36	49.3	253	2	Q8FHD2	Q8fhd2 escherichia
164	37	50.7	158	2	Q86D08	Q86d08 caenorhabdi	237	36	49.3	266	2	Q8EJ20	Q8ej20 shewanella
165	37	50.7	159	2	Q67VQ7	Q67vq7 oryza sativ	238	36	49.3	279	2	Q80VI5	Q80vi5 mus musculu
166	37	50.7	185	1	RRF_MYCLE	Q33046 mycobacteri	239	36	49.3	280	2	Q95698	Q95698 homo sapien
167	37	50.7	187	2	Q8HU07	Q8hu07 rhodobryum	240	36	49.3	280	2	Q9H160	Q9h160 homo sapien
168	37	50.7	187	2	Q68IK5	Q68ik5 helicobacte	241	36	49.3	281	2	Q8BGU8	Q8bgu8 m mus muscu
169	37	50.7	255	2	Q8MI30	Q8mi30 equus cabal	242	36	49.3	281	2	Q9ESK4	Q9esk4 mus musculu
170	37	50.7	324	2	O04083	O04083 arabidopsis	243	36	49.3	292	2	Q90Z34	Q90z34 brachydanio
171	37	50.7	330	2	Q82MF0	Q82mf0 streptomyce	244	36	49.3	306	2	Q7RIN6	Q7rin6 plasmodium
172	37	50.7	338	2	Q8SSM8	Q8ssm8 encephalito	245	36	49.3	328	2	Q98K20	Q98k20 rhizobium l
173	37	50.7	356	2	Q919T1	Q919t1 influenza a	246	36	49.3	329	1	PIN2_CAEEL	Q19157 caenorhabdi
174	37	50.7	383	2	Q6EDQ1	Q6edq1 hirundo pyr	247	36	49.3	340	2	Q8V256	Q8v256 helicoverpa
175	37	50.7	394	2	Q8EVB5	Q8evb5 mycoplasma	248	36	49.3	340	2	Q77LV8	Q77lv8 helicoverpa
176	37	50.7	410	2	Q45490	Q45490 caenorhabdi	249	36	49.3	340	2	Q9E219	Q9e219 helicoverpa
177	37	50.7	430	1	SYS_SYNY3	P73201 synechocyt	250	36	49.3	340	2	Q91FJ4	Q91fj4 helicoverpa

251	36	49.3	347	2	Q9XAN5	Q9xan5 streptomyce	324	36	49.3	822	1	MUS2_SYN3	P73625 synechocyst
252	36	49.3	401	2	Q6D454	Q6d454 erwinia car	325	36	49.3	897	2	Q93Y95	Q93y95 zea mays (m
253	36	49.3	402	2	Q93YD4	Q93yd4 nicotiana t	326	36	49.3	996	2	Q89ZU9	Q89zu9 bacteroides
254	36	49.3	402	2	Q93YB5	Q93yb5 nicotiana t	327	36	49.3	1063	2	Q61FK6	Q61fk6 plasmodium
255	36	49.3	402	2	Q948Y4	Q948y4 nicotiana t	328	36	49.3	1168	2	Q7RH99	Q7rh99 plasmodium
256	36	49.3	402	2	Q9XG74	Q9xg74 nicotiana t	329	36	49.3	1247	2	Q817Z6	Q817z6 toxoplasma
257	36	49.3	404	2	Q7N621	Q7n621 photorhabd	330	36	49.3	1363	2	Q81AM5	Q81am5 plasmodium
258	36	49.3	406	2	Q93YC1	Q93yc1 nicotiana t	331	36	49.3	1702	2	P97706	P97706 rattus norv
259	36	49.3	411	2	Q9Y8N4	Q9y8n4 aeropyrum p	332	36	49.3	1907	2	Q96L96	Q96l96 homo sapien
260	36	49.3	428	2	Q6GD81	Q6gd81 staphylococ	333	36	49.3	2209	2	Q9U0G6	Q9u0g6 plasmodium
261	36	49.3	435	2	Q01340	Q01340 rattus norv	334	36	49.3	3454	2	Q7Y1G0	Q7y1g0 oryza sativ
262	36	49.3	461	2	Q28377	Q28377 archaeoglob	335	36	49.3	5432	2	Q7RXP5	Q7rxp5 plasmodium
263	36	49.3	466	2	Q9S5X4	Q9s5x4 thermotoga	336	35.5	48.6	468	2	Q72WL4	Q72wl4 desulfovibr
264	36	49.3	470	2	Q9WYR5	Q9wyrs thermotoga	337	35.5	48.6	2112	2	Q9VEL9	Q9vel9 drosophila
265	36	49.3	481	2	Q7MXE2	Q7mxex porphyromon	338	35	47.9	55	2	Q6SA81	Q6sa81 salmonella
266	36	49.3	485	2	Q8CFK1	Q8cfk1 mus musculu	339	35	47.9	104	2	Q51757	Q51757 pseudomonas
267	36	49.3	488	2	Q75DX4	Q75dx4 ashbya gos	340	35	47.9	110	2	Q942Y1	Q942y1 oryza sativ
268	36	49.3	498	1	NMT2_HUMAN	Q60551 homo sapien	341	35	47.9	112	2	Q46132	Q46132 clostridium
269	36	49.3	505	2	Q89M54	Q89m54 bradyrhizob	342	35	47.9	119	2	Q89K90	Q89k90 bradyrhizob
270	36	49.3	505	2	Q89NR7	Q89nr7 rhizobium l	343	35	47.9	128	2	Q74FF9	Q74ff9 geobacter s
271	36	49.3	510	2	Q6W0Z3	Q6w0z3 methanococ	344	35	47.9	137	2	Q88JY2	Q88jy2 pseudomonas
272	36	49.3	514	2	Q8KZ28	Q8kz28 uncultured	345	35	47.9	140	2	Q6JLB3	Q6jlb3 franciella
273	36	49.3	523	2	Q6K632	Q6k632 oryza sativ	346	35	47.9	173	2	Q8EIH1	Q8eih1 shewanella
274	36	49.3	529	1	NMT2_MOUSE	Q70311 mus musculu	347	35	47.9	179	2	Q67BF6	Q67bf6 haemophilus
275	36	49.3	529	2	Q700Q7	Q700q7 rattus norv	348	35	47.9	179	2	Q7VN40	Q7vn40 haemophilus
276	36	49.3	530	2	Q6BQJ5	Q6bjj5 debaryomyce	349	35	47.9	187	1	AMPD_SALTY	P30013 salmonella
277	36	49.3	532	2	Q8KX52	Q8kx52 synechococ	350	35	47.9	187	2	Q8Z9F5	Q8z9f5 salmonella
278	36	49.3	537	2	Q6Z543	Q6z543 oryza sativ	351	35	47.9	190	2	Q96BF5	Q96bf5 homo sapien
279	36	49.3	601	2	Q75DW2	Q75dw2 ashbya gos	352	35	47.9	195	2	Q9VRN5	Q9vrn5 drosophila
280	36	49.3	609	2	Q81ER1	Q81er1 plasmodium	353	35	47.9	200	2	O02498	O02498 leishmania
281	36	49.3	620	2	Q6UJF6	Q6ujf6 drosophila	354	35	47.9	205	1	BASI_BOVIN	Q865r3 bos taurus
282	36	49.3	620	2	Q6UJF7	Q6ujf7 drosophila	355	35	47.9	217	2	Q6PVK2	Q6pvk2 pseudomonas
283	36	49.3	620	2	Q6UJF9	Q6ujf9 drosophila	356	35	47.9	221	2	Q9ALZ8	Q9alz8 fremyella d
284	36	49.3	621	2	Q6UJG2	Q6ujg2 drosophila	357	35	47.9	229	2	Q6M9T8	Q6m9t8 parachlamyd
285	36	49.3	621	2	Q9NAY2	Q9nay2 drosophila	358	35	47.9	237	2	Q6AME6	Q6ame6 desulfotale
286	36	49.3	621	2	Q9NGM3	Q9ngm3 drosophila	359	35	47.9	237	2	Q8EUM8	Q8eum8 mus musculu
287	36	49.3	630	2	Q6UJF8	Q6ujf8 drosophila	360	35	47.9	239	2	O44730	O44730 caenorhabdi
288	36	49.3	630	2	Q6UJG0	Q6ujg0 drosophila	361	35	47.9	252	2	Q9LIQ7	Q9liq7 streptomyce
289	36	49.3	630	2	Q9NAY3	Q9nay3 drosophila	362	35	47.9	255	2	Q86KQ6	Q86kq6 dictyosteli
290	36	49.3	630	2	Q9NAY4	Q9nay4 drosophila	363	35	47.9	256	2	Q86KQ8	Q86kq8 dictyosteli
291	36	49.3	635	1	HS68_DROME	Q97125 drosophila	364	35	47.9	259	2	Q6LGM9	Q6lqw9 photobacter
292	36	49.3	639	2	Q7QUK1	Q7quk1 giardia lam	365	35	47.9	262	2	Q8Y1G0	Q8ylg0 ralatonia s
293	36	49.3	644	1	SYR_AERPE	Q9y339 aeropyrum p	366	35	47.9	283	2	Q9W057	Q9w057 drosophila
294	36	49.3	687	2	Q69DK8	Q69dk8 sus scrofa	367	35	47.9	284	1	LPPI_HUMAN	O14494 homo sapien
295	36	49.3	707	2	Q9CD82	Q9cd82 mycobacteri	368	35	47.9	292	2	Q7NFX3	Q7nfx3 gloeobacter
296	36	49.3	711	2	O81292	O8i292 plasmodium	369	35	47.9	293	2	Q8RYM8	Q8rym8 oryza sativ
297	36	49.3	732	2	Q86UF2	Q86uf2 homo sapien	370	35	47.9	320	2	Q702C4	Q702c4 uncultured
298	36	49.3	747	2	Q8QPW0	Q8qpw0 influenza b	371	35	47.9	321	2	Q9GRJ9	Q9grj9 leishmania
299	36	49.3	752	1	RRP1_INBAC	P13871 influenza b	372	35	47.9	321	2	Q8RYK1	Q8ryk1 oryza sativ
300	36	49.3	752	1	RRP1_INBAD	P13872 influenza b	373	35	47.9	341	2	Q9HLX8	Q9hlx8 thermoplaem
301	36	49.3	752	1	RRP1_INBLE	P07832 influenza b	374	35	47.9	342	2	Q8LCP0	Q8lcp0 arabidopsis
302	36	49.3	752	1	RRP1_INBP9	O36430 influenza b	375	35	47.9	342	2	Q9SND6	Q9snd6 arabidopsis
303	36	49.3	752	2	Q681P2	Q681p2 influenza b	376	35	47.9	346	2	Q9SUZ8	Q9suz8 arabidopsis
304	36	49.3	752	2	Q77SM9	Q77sm9 influenza b	377	35	47.9	347	2	Q9YCM3	Q9ycm3 aeropyrum p
305	36	49.3	752	2	Q77SN0	Q77sn0 influenza b	378	35	47.9	362	2	Q6IL91	Q6il91 drosophila
306	36	49.3	752	2	Q77SN1	Q77sn1 influenza b	379	35	47.9	377	2	Q88DG5	Q88dg5 pseudomonas
307	36	49.3	752	2	Q77SN2	Q77sn2 influenza b	380	35	47.9	382	2	Q7UJS6	Q7ujs6 rhodopirell
308	36	49.3	752	2	Q80DM8	Q80dm8 influenza b	381	35	47.9	385	1	BASI_HUMAN	P35613 h basigin p
309	36	49.3	752	2	Q80DN5	Q80dn5 influenza b	382	35	47.9	392	2	Q94B20	Q94b20 arabidopsis
310	36	49.3	752	2	Q91MP8	Q91mp8 influenza b	383	35	47.9	392	2	Q9SA96	Q9sa96 arabidopsis
311	36	49.3	752	2	Q9PWU1	Q9pwu1 influenza b	384	35	47.9	401	2	Q82N33	Q82n33 streptomyce
312	36	49.3	752	2	Q9QLJ8	Q9qlj8 influenza b	385	35	47.9	415	2	Q8LB99	Q8lb99 arabidopsis
313	36	49.3	752	2	Q9QLJ9	Q9qlj9 influenza b	386	35	47.9	415	2	Q9LJ68	Q9lj68 arabidopsis
314	36	49.3	752	2	Q9QLK0	Q9qlk0 influenza b	387	35	47.9	419	2	Q7RLU9	Q7rlu9 plasmodium
315	36	49.3	752	2	Q9QLK1	Q9qlk1 influenza b	388	35	47.9	419	2	Q73UF6	Q73uf6 mycobacteri
316	36	49.3	752	2	Q9QLK2	Q9qlk2 influenza b	389	35	47.9	422	1	PGL1_ARATH	P49062 arabidopsis
317	36	49.3	752	2	Q9QLK3	Q9qlk3 influenza b	390	35	47.9	422	2	Q8K426	Q8k426 streptomyce
318	36	49.3	752	2	Q9QLK4	Q9qlk4 influenza b	391	35	47.9	431	2	Q8PDE1	Q8pde1 xanthomonas
319	36	49.3	752	2	Q9QLK5	Q9qlk5 influenza b	392	35	47.9	435	2	O16968	O16968 caenorhabdi
320	36	49.3	752	2	Q9QLK6	Q9qlk6 influenza b	393	35	47.9	439	1	L1A3_HUMAN	Q8n6c8 homo sapien
321	36	49.3	752	2	Q9QLK7	Q9qlk7 influenza b	394	35	47.9	447	2	Q9F0G1	Q9f0g1 streptococc
322	36	49.3	752	2	Q9QLK8	Q9qlk8 influenza b	395	35	47.9	447	2	Q9FDM8	Q9fdm8 streptococc
323	36	49.3	752	2	Q9QLK9	Q9qlk9 influenza b	396	35	47.9	447	2	Q9RPL5	Q9rpl5 streptococc

397	35	47.9	453	2	P92013	P92013 caenorhabdi	470	34	46.6	122	2	Q40738	Q40738 oryza satia	
398	35	47.9	471	1	Y1R1 DROME	P16424 drosophila	471	34	46.6	124	2	Q91TY9	Q91ty9 influenza a	
399	35	47.9	482	2	Q23260	Q23260 arabidopsis	472	34	46.6	129	2	Q91TZ1	Q91tz1 influenza a	
400	35	47.9	488	2	Q8K0T4	Q8K0T4 mus musculu	473	34	46.6	137	2	Q8NAF9	Q8naf9 homo sapien	
401	35	47.9	502	1	LAS1 YEAST	P36146 saccharomyc	474	34	46.6	146	2	Q9CTI3	Q9cti3 mus musculu	
402	35	47.9	502	2	Q6GR55	Q66r55 saccharomyc	475	34	46.6	159	2	Q8N8D2	Q8n8d2 homo sapien	
403	35	47.9	503	2	Q6BK51	Q6bk51 debaryomyc	476	34	46.6	164	2	Q8HRE9	Q8hre9 bryum gemmi	
404	35	47.9	508	2	Q7UXZ0	Q7uxz0 rhodopirell	477	34	46.6	186	2	Q8HRF4	Q8hrf4 anomobryum	
405	35	47.9	521	2	Q7S3M1	Q7s3m1 neurospora	478	34	46.6	187	2	Q8MG46	Q8mg46 rosulabryum	
406	35	47.9	521	2	Q8AR16	Q8ar16 arabidopsis	479	34	46.6	188	2	Q63005	Q63005 bryum alpin	
407	35	47.9	546	1	PTR2 MOUSE	Q91v95 mus musculu	480	34	46.6	188	2	Q9AY14	Q9ay14 physcometri	
408	35	47.9	551	2	Q8T294	Q8t294 dictyosteli	481	34	46.6	188	2	Q91RL9	Q91rl9 arabidopsis	
409	35	47.9	562	2	Q6ZMG3	Q6zmg3 homo sapien	482	34	46.6	188	2	Q6JKP4	Q6jkp4 influenza a	
410	35	47.9	569	2	Q84H24	Q84h24 pseudomonas	483	34	46.6	188	2	Q6JKP5	Q6jkp5 influenza a	
411	35	47.9	575	2	Q8WTU2	Q8wtu2 homo sapien	484	34	46.6	188	2	Q6JKP6	Q6jkp6 influenza a	
412	35	47.9	575	2	Q9G645	Q9g645 homo sapien	485	34	46.6	188	2	Q6JKP7	Q6jkp7 influenza a	
413	35	47.9	581	2	Q89244	Q89244 marek's dis	486	34	46.6	188	2	Q63006	Q63006 bryum caesp	
414	35	47.9	581	2	Q782P8	Q782p8 gallid herp	487	34	46.6	189	2	Q63011	Q63011 rhodobryum w	
415	35	47.9	622	2	Q653D2	Q653d2 oryza sativ	488	34	46.6	189	2	Q6JE74	Q6je74 sphlachnum w	
416	35	47.9	626	2	Q811Z9	Q811z9 plasmodium	489	34	46.6	190	2	Q62998	Q62998 funaria hyg	
417	35	47.9	627	2	Q7RWT2	Q7rwt2 neurospora	490	34	46.6	190	2	Q63009	Q63009 bryum steno	
418	35	47.9	628	2	Q6A849	Q6a849 propionibac	491	34	46.6	190	2	Q8HU00	Q8hu00 bryum palie	
419	35	47.9	631	2	Q8SVW5	Q8svw5 encephalito	492	34	46.6	190	2	Q8MCT1	Q8mct1 haplodontiu	
420	35	47.9	638	2	Q9GLR0	Q9glr0 bos taurus	493	34	46.6	190	2	Q6K3W6	Q6k3w6 oryza sativ	
421	35	47.9	656	2	Q9M092	Q9m092 arabidopsis	494	34	46.6	190	2	Q9AY13	Q9ay13 aphanorrrheg	
422	35	47.9	662	1	I12R HUMAN	P42701 homo sapien	495	34	46.6	191	2	Q9AY15	Q9ay15 physcometri	
423	35	47.9	727	2	Q6RKB2	Q6rkb2 rattus norv	496	34	46.6	191	2	Q6W760	Q6w760 influenza a	
424	35	47.9	735	2	Q9N8A6	Q9n8a6 trypanosoma	497	34	46.6	191	2	Q6W761	Q6w761 influenza a	
425	35	47.9	739	2	Q9Q0S6	Q9q0s6 influenza a	498	34	46.6	191	2	Q8BCD5	Q8bcd5 influenza a	
426	35	47.9	757	2	Q6LTA6	Q6lta6 photobacter	499	34	46.6	191	2	Q8BCD6	Q8bcd6 influenza a	
427	35	47.9	782	2	Q8PP65	Q8pp65 xanthomonas	500	34	46.6	191	2	Q8BCE4	Q8bce4 influenza a	
428	35	47.9	837	1	NCM2 HUMAN	O15394 homo sapien								
429	35	47.9	837	1	NCM2 MOUSE	Q35136 mus musculu								
430	35	47.9	837	2	Q727F2	Q727f2 homo sapien								
431	35	47.9	837	2	Q6RKB3	Q6rkb3 rattus norv								
432	35	47.9	849	2	Q815V7	Q815v7 plasmodium								
433	35	47.9	958	2	Q7NNV9	Q7nnv9 chromobacte								
434	35	47.9	1019	2	Q6CV73	Q6cv73 kluyveromyc								
435	35	47.9	1036	2	Q6NVF2	Q6nvf2 mus musculu								
436	35	47.9	1065	2	Q6F2X0	Q6f2x0 oryza sativ								
437	35	47.9	1074	1	PLD1 MOUSE	Q92280 mus musculu								
438	35	47.9	1074	1	FLD1 RAT	P70496 rattus norv								
439	35	47.9	1137	2	Q8T9H7	Q8t9h7 drosophila								
440	35	47.9	1137	2	Q9VXF4	Q9vxf4 drosophila								
441	35	47.9	1181	2	Q96JV0	Q96jv0 homo sapien								
442	35	47.9	1234	2	Q7P1Q7	Q7p1q7 anopheles g								
443	35	47.9	1322	2	Q7FNR7	Q7fnr7 anopheles g								
444	35	47.9	1322	2	Q9NAT0	Q9nat0 anopheles g								
445	35	47.9	1322	2	Q9NJS5	Q9njs5 anopheles g								
446	35	47.9	1323	2	O13686	O13686 schizosacch								
447	35	47.9	1402	2	Q7RBO6	Q7rbq6 plasmodium								
448	35	47.9	1412	1	Z462 HUMAN	Q96jm2 homo sapien								
449	35	47.9	1499	2	Q9R839	Q9r839 mycosphaere								
450	35	47.9	1559	2	Q81DC3	Q81dc3 plasmodium								
451	35	47.9	1779	2	Q9BPP6	Q9bpb6 samia cynth								
452	35	47.9	2153	2	Q9R7375	Q9r7375 strongyloce								
453	35	47.9	2403	2	Q68CP0	Q68cp0 homo sapien								
454	35	47.9	2506	2	Q63HJ5	Q63hj5 homo sapien								
455	35	47.9	3115	2	Q9VQV1	Q9vqv1 drosophila								
456	34.5	47.3	325	2	Q9P225	Q9p225 campylobact								
457	34.5	47.3	555	2	Q8H5J9	Q8h5j9 oryza sativ								
458	34	46.6	67	2	Q82548	Q82548 influenza a								
459	34	46.6	72	2	Q26942	Q26942 tribolium c								
460	34	46.6	77	2	Q7S1T4	Q7slt4 neurospora								
461	34	46.6	78	2	Q82547	Q82547 influenza a								
462	34	46.6	85	2	Q7XT60	Q7xt60 oryza sativ								
463	34	46.6	87	2	Q9P224	Q9p224 homo sapien								
464	34	46.6	90	2	Q621J9	Q621j9 oryza sativ								
465	34	46.6	92	2	Q9EYK7	Q9eyk7 lactobacill								
466	34	46.6	105	2	Q7XAW8	Q7xaw8 brassica ra								
467	34	46.6	105	2	Q7YAX1	Q7yax1 brassica ra								
468	34	46.6	112	2	Q50739	Q50739 borrelia bu								
469	34	46.6	116	2	Q5RYF3	Q5ryf3 deinococcus								

ALIGNMENTS													
RESULT 1	LPP3_HUMAN	STANDARD;	PRT;	311 AA.									
ID	LPP3_HUMAN	STANDARD;	PRT;	311 AA.									
AC	O14495; Q96GW0; Q99782;												
DT	29-MAR-2004 (Rel. 43, Created)												
DT	29-MAR-2004 (Rel. 43, Last sequence update)												
DT	05-JUL-2004 (Rel. 44, Last annotation update)												
DE	Lipid phosphate phosphohydrolase 3 (EC 3.1.3.4) (Phosphatidic acid												
DE	phosphatase 2b) (Phosphatidate phosphohydrolase type 2b) (PAP2b) (PAP-												
DE	2b) (PAP2-beta) (Vascular endothelial growth factor and type 1												
DE	collagen inducible protein) (VCIP).												
GN	Name=PPAP2B; Synonym=PP3;												
OS	Homo sapiens (Human).												
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;												
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.												
OX	NCBI_TaxID=9606;												
RP	[1]												
RP	SEQUENCE FROM N.A., CHARACTERIZATION, AND INDUCTION.												
RX	MEDLINE=97450990; PubMed=9305923; DOI=10.1074/jbc.272.39.24572;												
RA	Kai M., Wada I., Imai S.-I., Sakane F., Kanoh H.;												
RT	"Cloning and characterization of two human isozymes of Mg2+-												
RT	independent phosphatidic acid phosphatase.";												
RL	J. Biol. Chem. 272:24572-24578(1997).												
RN	[2]												
RP	SEQUENCE FROM N.A., AND CHARACTERIZATION.												
RX	MEDLINE=98371049; PubMed=9705349; DOI=10.1074/jbc.273.34.22059;												
RA	Roberts R., Sciorra V.A., Morris A.J.;												
RT	"Human type 2 phosphatidic acid phosphohydrolases. Substrate												
RT	specificity of the type 2a, 2b, and 2c enzymes and cell surface												
RT	activity of the 2a isoform.";												
RL	J. Biol. Chem. 273:22059-22067(1998).												
RN	[3]												
RP	SEQUENCE FROM N.A.												
RX	PubMed=12660161; DOI=10.1093/emboj/cdg165;												
RA	Humzoe J.O., Feng S., Thakker G.D., Yang J., Hong J., Wary K.K.;												
RT	"Regulation of cell-cell interactions by phosphatidic acid phosphatase												
RT	2b/VCIP.";												

ALIGNMENTS

RESULT 1	LP3_HUMAN	STANDARD;	PRT;	311 AA.
AC	O14495; Q96GW0; Q99782;			
DT	29-MAR-2004 (Rel. 43, Last sequence update)			
DT	29-MAR-2004 (Rel. 43, Last sequence update)			
DE	Lipid phosphate phosphohydrolase 3 (EC 3.1.3.4) (Phosphatidic acid			
DE	phosphatase 2b) (Phosphatidate phosphohydrolase type 2b) (PAP-2b)			
DE	collagen inducible protein) (VCIP).			
OS	Homo sapiens (Human).			
GN	Name=PPAP2B; Synonyms=LP3;			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A., CHARACTERIZATION, AND INDUCTION.			
RX	MEDLINE=97450390; PubMed=9305923; DOI=10.1074/jbc.272.39.24572;			
RA	Kai M., Wada I., Imai S.-I., Sakane F., Kanch H.;			
RT	"Cloning and characterization of two human isozymes of Mg2+-			
RT	independent phosphatidic acid phosphatase.";			
RL	J. Biol. Chem. 272:24572-24578 (1997).			
RN	[2]			
RP	SEQUENCE FROM N.A., AND CHARACTERIZATION.			
RX	MEDLINE=98371049; PubMed=9705349; DOI=10.1074/jbc.273.34.22059;			
RA	Roberts R., Sciorra V.A., Morris A.J.;			
RT	"Human type 2 phosphatidic acid phosphohydrolases. Substrate			
RT	specificity of the type 2a, 2b, and 2c enzymes and cell surface			
RT	activity of the 2a isoform.";			
RL	J. Biol. Chem. 273:22059-22067 (1998).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	PubMed=12660161; DOI=10.1093/emboj/cdg165;			
RA	Huntsue J.O., Feng S., Thakker G.D., Yang J., Hong J., Wary K.K.;			
RT	"Regulation of cell-cell interactions by phosphatidic acid phosphatase			
RT	2b/VCIP.";			

EMBO J. 22:1539-1554 (2003).
 [4]
 SEQUENCE FROM N.A.
 Leung D.W., Tompkins C.K.;
 RT "Molecular cloning of and expression of an isoform of human
 RT phosphatidic acid phosphatase cDNA";
 RT Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 [5]
 SEQUENCE FROM N.A.
 MEDLINE=97264341; PubMed=9110174;
 RX Yu W., Anderson B., Worley K.C., Muzny D.M., Ding Y., Liu W.,
 RA Ricafrente J.Y., Wentland M.A., Lennon G., Gibbs R.A.;
 RA "Large-scale concatenation cDNA sequencing";
 RL Genome Res. 7:353-358 (1997).
 [6]
 SEQUENCE FROM N.A.
 TISSUE=Placenta;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Whiting M., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
 RA Schnarch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 [7]
 SUBUNIT.
 RX PubMed=14725715; DOI=10.1186/1471-2091-5-2;
 RA Burnett C., Makridou P., Hewlett L., Howard K.;
 RA "Lipid phosphate phosphatases dimerise, but this interaction is not
 RT required for in vivo activity";
 RL BMC Biochem. 5:2-2 (2004).
 CC -!- FUNCTION: Catalyzes the conversion of phosphatidic acid (PA) to
 CC diacylglycerol (DG). In addition it hydrolyzes lysophosphatidic
 CC acid (LPA), ceramide-1-phosphate (C-1-P) and sphingosine-1-
 CC phosphate (S-1-P). The relative catalytic efficiency is LPA = PA >
 CC C-1-P > S-1-P. May be involved in cell adhesion and in cell-cell
 CC interactions.
 CC -!- CATALYTIC ACTIVITY: A 3-sn-phosphatidate + H(2O) = a 1,2-diacyl-
 CC sn-glycerol + phosphate.
 CC -!- ENZYME REGULATION: Inhibited by sphingosine, zinc ions and
 CC propanolol. Not inhibited by N-ethylmaleimide treatment.
 CC -!- SUBUNIT: Homodimer. This complex seems not to be involved in
 CC substrate recognition, it may confer only structural or functional
 CC stability.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Post-Golgi and
 CC plasma membrane localization.
 CC -!- TISSUE SPECIFICITY: Ubiquitously expressed. Highly expressed in
 CC heart and placenta.
 CC -!- INDUCTION: By epidermal growth factor (EGF), vascular endothelial
 CC growth factor (VEGF), basic fibroblast growth factor (bFGF) and
 CC phorbol myristate acetate (PMA).
 CC -!- PTM: N-glycosylated. Contains high-mannose oligosaccharides.
 CC -!- SIMILARITY: Belongs to the PA-phosphatase related phosphoesterase
 CC family.
 CC -!- CAUTION: Ref.5 sequence differs from that shown due to a
 CC frameshift in position 225.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 DR EMBL; AB000889; BAA22594.1; -;
 DR EMBL; AF017786; AAC63433.1; -;
 DR EMBL; AF408083; AAO84481.1; -;
 DR EMBL; AF043329; AAO02271.1; -;
 DR EMBL; U79294; AAB50222.1; ALT FRAME.
 DR EMBL; BC009196; AAH09196.1; -;
 DR Genew; HGNC:9229; PPAP2B.
 DR H-InvDB; HIX0000628; -;
 DR Reactome; O14495; -;
 DR MIM; 607125; -;
 DR GO; GO:0016020; C:membrane; TAS.
 DR GO; GO:0004721; P:phosphoprotein phosphatase activity; TAS.
 DR GO; GO:0008151; P:cell growth and/or maintenance; TAS.
 DR GO; GO:0008354; P:germ-cell migration; TAS.
 DR InterPro; IPR0008934; ACPase_VanPerase.
 DR InterPro; IPR000326; Peptase_PA_PTP.
 DR Pfam; PF01569; PAP2; 1.
 DR SMART; SM00014; acidPpc; 1.
 KW Glycoprotein; Hydrolase; Transmembrane.
 FT DOMAIN 1 33 Cytoplasmic (Potential).
 FT TRANSMEM 34 54 Potential.
 FT DOMAIN 55 85 Lumenal (Potential).
 FT TRANSMEM 86 106 Potential.
 FT DOMAIN 107 122 Cytoplasmic (Potential).
 FT TRANSMEM 123 143 Potential.
 FT DOMAIN 144 193 Lumenal (Potential).
 FT TRANSMEM 194 214 Potential.
 FT DOMAIN 215 227 Cytoplasmic (Potential).
 FT TRANSMEM 228 248 Potential.
 FT DOMAIN 249 257 Lumenal (Potential).
 FT TRANSMEM 258 278 Potential.
 FT DOMAIN 279 311 Cytoplasmic (Potential).
 FT CARBOHYD 170 170 N-linked (GlcNAc...) (Potential).
 FT CONFLICT 282 282 T -> M (in Ref. 6).
 SQ SEQUENCE 311 AA; 35116 MW; CB3F60189044DA31 CRC64;
 Query Match 100.0%; Score 73; DB 1; Length 311;
 Best Local Similarity 100.0%; Pred. No. 8.6e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NYRCRGDDSKVQE 13
 DB 178 NYRCRGDDSKVQE 190
 RESULT 2
 LPP3 MOUSE
 ID LPP3 MOUSE STANDARD; PRT; 312 AA.
 AC Q99YF8; QB8TB7;
 DT 29-MAR-2004 (Rel. 43, Created)
 DT 29-MAR-2004 (Rel. 43, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Lipid phosphate phosphohydrolase 3 (EC 3.1.3.4) (Phosphatidic acid
 DE phosphatase 2b) (Phosphatidate phosphohydrolase type 2b) (PAP2b) (PAP-
 DE 2b) (PAP2-beta).
 GN Name=Pap2b; Synonyms=Lpp3;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryo;
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojibori T.,

RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustinchik S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Malais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Perte G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada M.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Inotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sakaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedon T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skaleka U., Smailus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP FUNCTION, AND DISEASE.
RX PubMed=12925589; DOI=10.1242/dev.00635;
RA Escalante-Alcalde D., Hernandez L., Le Stunff H., Maeda R., Lee H.-S.,
RA Cheng G. Jr., Sciorra V.A., Daar I., Spiegel S., Morris A.J.,
RA Stewart C.L.;
RT "The lipid phosphatase LPP3 regulates extra-embryonic vasculogenesis
RT and axis patterning";
RL Development 130:4623-4637 (2003).
CC -!- FUNCTION: Catalyzes the conversion of phosphatidic acid (PA) to
CC diacylglycerol (DG). In addition it hydrolyzes lysophosphatidic
CC acid (LPA), ceramide-1-phosphate (C-1-P) and sphingosine-1-
CC phosphate (S-1-P) (By similarity). Essential to the formation of
CC the chorio-allantoic placenta and extra-embryonic vasculature.
CC Also mediates gastrulation and axis formation, probably by
CC regulating the Wnt signaling pathway.
CC -!- CATALYTIC ACTIVITY: A 3-sn-phosphatidate + H(2)O = 1,2-diacyl-
CC sn-glycerol + phosphate.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC reticulum associated (By similarity).
CC -!- DISEASE: Ppap2b deficient embryos fail to form a chorio-allantoic
CC placenta and yolk sac vasculature. A subset of embryos also show a
CC shortening of the anterior-posterior axis and frequent duplication

CC of axial structures. Loss of Ppap2b results in a marked increase
CC in beta-catenin-mediated T-cell factor (TCF) transcription.
CC -!- SIMILARITY: Belongs to the PA-phosphatase related phosphoesterase
CC family.
CC -!- CAUTION: Ref.1 sequence differs from that shown due to a
CC frameshift in position 226.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AK011276; BAC35327.1; ALT_FRAME.
CC EMBL; BC005558; AAH05558.1; -.
DR MGD; MGI:1915166; Ppap2b.
DR GO; GO:0042577; F:lipid phosphatase activity; IMP.
DR GO; GO:0001568; P:blood vessel development; IMP.
DR GO; GO:0010003; P:gastrulation (sensu Mammalia); IMP.
DR GO; GO:0006644; P:phospholipid metabolism; IMP.
DR GO; GO:0030111; P:regulation of wnt receptor signaling pathway; IDA.
DR InterPro; IPR008934; AcPase_VanPerase.
DR InterPro; IPR000326; Pesterase_PA_ptp.
DR Pfam; PF01569; PAP2; 1.
DR SMART; SM00014; acidPPc; 1.
KW Developmental protein; Glycoprotein; Hydrolase; Transmembrane.
FT DOMAIN 1 33 Cytoplasmic (Potential).
FT TRANSMEM 34 54 Potential.
FT DOMAIN 55 85 Luminal (Potential).
FT TRANSMEM 86 106 Potential.
FT DOMAIN 107 123 Cytoplasmic (Potential).
FT TRANSMEM 124 144 Potential.
FT DOMAIN 145 194 Luminal (Potential).
FT TRANSMEM 195 215 Potential.
FT DOMAIN 216 226 Cytoplasmic (Potential).
FT TRANSMEM 227 247 Potential.
FT DOMAIN 248 258 Luminal (Potential).
FT TRANSMEM 259 279 Potential.
FT DOMAIN 280 312 Cytoplasmic (Potential).
FT CARBOHYD 171 171 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 312 AA; 35216 MW; D782986E04B57D7D CRC64;
Query Match 94.5%; Score 69; DB 1; Length 312;
Best Local Similarity 92.3%; Pred. No. 0.00046;
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 NYRCRGDDSKVQE 13
Db 179 NYRCRGDDSKVQE 191
RESULT 3
LPP3_RAT
ID_LPP3_RAT STANDARD; PRT; 312 AA.
AC P97544;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Lipid phosphate phosphohydrolase 3 (SC 3.1.3.4) (Phosphatidic acid
DE phosphate 2b) (Phosphatidate phosphohydrolase type 2b) (PAP2b) (PAP-
DE 2b) (PAP2-beta) (Differentially expressed in rat intestine 42)
DE (Dri42).
DE Name=Ppap2b; Synonyms=Lpp3;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., TOPOLOGY, AND N-GLYCOSYLATION.
RC STRAIN=Wistar; TISSUE=Small intestine;
RX MEDLINE=97094703; PubMed=8939937; DOI=10.1074/jbc.271.47.29928;

RA Barila D., Plateroti M., Nobili F., Muda A.O., Xie Y., Morimoto T.,
 RA Perozzi G.;
 RT "The Dri 42 gene, whose expression is upregulated during epithelial
 RT differentiation, encodes a novel ER resident transmembrane protein.";
 RL J. Biol. Chem. 271:29928-29936(1996).
 RN [2]
 RP TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
 RX PubMed=8055940;
 RA Barila D., Murgia C., Nobili F., Gaetani S., Perozzi G.;
 RT "Subtractive hybridization cloning of novel genes differentially
 RT expressed during intestinal development.";
 RL Eur. J. Biochem. 223:701-709(1994).
 RN [3]
 RP TISSUE SPECIFICITY.
 RC STRAIN=Sprague-Dawley;
 RX MEDLINE=2155999; PubMed=11704545;
 RA Nanjundan M., Possmayer F.;
 RT "Molecular cloning and expression of pulmonary lipid phosphate
 RT phosphohydrolases";
 RL Am. J. Physiol. 281:L1484-L1493(2001).
 CC -1- FUNCTION: Catalyzes the conversion of phosphatidic acid (PA) to
 CC diacylglycerol (DG). In addition it hydrolyzes lysophosphatidic
 CC acid (LPA), ceramide-1-phosphate (C-1-P) and sphingosine-1-
 CC phosphate (S-1-P) (By similarity). Involved in the regulation of
 CC epithelial differentiation.
 CC -1- CATALYTIC ACTIVITY: A 3-sn-phosphatidate + H(2)O = a 1,2-diacyl-
 CC sn-glycerol + phosphate.
 CC -1- SUBUNIT: Homodimer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
 CC reticulum associated.
 CC -1- TISSUE SPECIFICITY: Detected in epithelial cells of intestinal
 CC mucosa, lung, liver and brain.
 CC -1- DEVELOPMENTAL STAGE: Expression is increased during epithelial
 CC differentiation in intestinal mucosa as well as in kidney, liver
 CC and lung.
 CC -1- PTM: N-glycosylated.
 CC -1- SIMILARITY: Belongs to the PA-phosphatase related phosphoesterase
 CC family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Y07783; CAA69106.1; -;
 DR InterPro: IPR008934; ACPase_VanPase.
 DR InterPro: IPR000326; Pesterase_PA_PTP.
 DR Pfam: PF01569; PAP2; 1.
 DR SMART: SM00014; acidPPC; 1.
 KW Endoplasmic reticulum; Glycoprotein; Hydrolase; Transmembrane.
 FT DOMAIN 1 33 Cytoplasmic (Potential).
 FT TRANSMEM 34 54 Potential.
 FT DOMAIN 55 85 Luminal (Potential).
 FT TRANSMEM 86 106 Potential.
 FT DOMAIN 107 123 Cytoplasmic (Potential).
 FT TRANSMEM 124 144 Potential.
 FT DOMAIN 145 194 Luminal (Potential).
 FT TRANSMEM 195 215 Potential.
 FT DOMAIN 216 226 Cytoplasmic (Potential).
 FT TRANSMEM 227 247 Potential.
 FT DOMAIN 248 258 Luminal (Potential).
 FT TRANSMEM 259 279 Potential.
 FT DOMAIN 280 312 Cytoplasmic (Potential).
 FT CARBOHYD 171 171 N-linked (GLCNAC...) (Potential).
 SQ SEQUENCE 312 AA; 35318 MW; 9B47FD321DB0419 CRC64;
 Query Match 94.5%; Score 69; DB 1; Length 312;
 Best Local Similarity 92.3%; Pred. No. 0.00046;
 Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYRCRGGDSKVQ 13
 DB 179 NYRCRGGDSKVQ 191
 RESULT 4
 Q6IMX4 PRELIMINARY; PRT; 312 AA.
 ID Q6IMX4;
 AC Q6IMX4;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE ER transmembrane protein Dri 42.
 GN Name=Dri42;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Heart;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Heart;
 RA Strausberg R.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC072544; AAH72544.1; -;
 DR GO: GO:0016021; C:integral to membrane; IEA.
 DR InterPro: IPR008934; ACPase_VanPase.
 DR InterPro: IPR000326; Pesterase_PA_PTP.
 DR Pfam: PF01569; PAP2; 1.
 DR SMART: SM00014; acidPPC; 1.
 KW Transmembrane.
 SQ SEQUENCE 312 AA; 35233 MW; CDA54495C087D37D CRC64;
 Query Match 94.5%; Score 69; DB 2; Length 312;
 Best Local Similarity 92.3%; Pred. No. 0.00046;
 Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYRCRGGDSKVQ 13
 DB 179 NYRCRGGDSKVQ 191
 RESULT 5
 Q68F34 PRELIMINARY; PRT; 307 AA.
 ID Q68F34;
 AC Q68F34;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE MGC81884 protein.

```

RN RP SEQUENCE FROM N.A.
RC STRAIN=k10;
RL Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RA Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
RR EMBL; AE017242; AS06848.1; -.
DR EMBL; AE017242; AS06848.1; -.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR003673; CAIB BAIF.
DR pfam; PF02515; CoA transf 3; 1.
DR Complete proteome; Hypothetical protein.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 776 AA; 82189 MW; 38A4528FCSB0B4BB CRC64;

Query Match 63.0%; Score 46; DB 2; Length 776;
Best Local Similarity 80.0%; Pred. No. 18;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 YRCRGDDSKV 11
DB 615 YRCRGDDAW 624
|||||:|

RESULT 7
Q95L84 PRELIMINARY; PRT; 167 AA.
ID Q95L84
AC Q95L84;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Phosphatidic acid phosphatase type 2A (Fragment).
DS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]

SEQUENCE FROM N.A.
RP MEDLINE=22306422; PubMed=12388084;
RA Wang D.-A., Du H., Jaggar J.H., Brindley D.N., Tigyi G.J.,
RA Watsky M.A.;
RT "Injury-elicited differential transcriptional regulation of
RT phospholipid growth factor receptors in the cornea.";
RL Am. J. Physiol. Cell Physiol. 283:C1646-C1654(2002).
DR EMBL; AF404277; AAL01884.1; -.
DR InterPro; IPR008934; AcPase VanPerase.
DR InterPro; IPR000326; Pesterase_PA_PTP.
DR pfam; PF01569; PAP2; 1.
DR SMART; SM00014; acidPPc; 1.
FT NON_TER 1
FT NON_TER 167
SQ SEQUENCE 167 AA; 18922 MW; 5185AECB0C8E8FB CRC64;

Query Match 60.3%; Score 44; DB 2; Length 167;
Best Local Similarity 61.5%; Pred. No. 7.8;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 NYVCRGDDSKVOE 13
DB 127 NYVCRGNAQKVXE 139
|||||:|

RESULT 8
Y787 METJA
ID Y787 METJA STANDARD; PRT; 504 AA.
AC Q58157;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hypothetical protein MJ0787.
GN OrderedLocusNames=MJ0787;
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]

```

DR PROSITE; P550075; ACP DOMAIN; 1.
 SQ SEQUENCE 705 AA; 75159 MW; 9F9D4F500690BFSD CRC64;

Query Match 60.3%; Score 44; DB 2; Length 705;
 Best Local Similarity 70.0%; Pred. No. 37;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps

QY 2 YKRCGDGDSKV 11
 |||||
 Db 383 YKRCADDTTEV 392

RESULT 10
 P96283 PRELIMINARY; PRT; 705 AA.

ID P96283 Q7D6D9;
 AC P96283; 07D6D9; 03, Created)
 DT 01-MAY-1997 (TReMBLrel. 03, Last annotation update)
 DT 01-MAY-1997 (TReMBLrel. 03, Last annotation update)
 DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
 DE PROBABLE FATTY-ACID-CoA LIASE FADD22 (FATTY-ACID-CoA SYNTHETASE)
 DE (FATTY-ACID-CoA SYNTHASE) (EC 6.2.1.1-) (Substrate--CoA ligase).
 GN Name-fadd22; OrderedLocusNames=MT3021, Rv2948c;
 OS Mycobacterium tuberculosis
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 NCBI_TaxID=1773;
 RX [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=H37Rv;
 RP MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
 RX Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
 RA Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,
 RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
 RA Connor R., Davies R.M., Devlin K., Peltwell T., Gentles S., Hamlin N.,
 RA Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
 RA Murphy J.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
 RA Rogers J., Rutter S., Seeger K., Skelton S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RL complete genome sequence.";
 RL Nature 393:537-544(1998).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RP MEDLINE=22206494; PubMed=12218036;
 RX DOI=10.1126/JB.184.19.5479-5490.2002;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
 RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
 RA Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,
 RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
 RA Fraser C.M.;
 RA "Whole-genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains.";
 RT J. Bacteriol. 184:5479-5490(2002).
 RC -1- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
 CC family.
 DR EMBL; BX842581; CAB06101.1; -;
 DR EMBL; AE000516; AAK47347.1; -;
 DR FIR; A70669; A70669.
 DR HSSP; P08659; 1LCI.
 DR TIGR; MT3021; -;
 DR TubercuList; Rv2948c; -;
 DR GO; GO:0003824; F:catalytic activity; IEA.
 DR GO; GO:0048037; F:cofactor binding; IEA.
 DR GO; GO:0016874; F:ligase activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR009081; ACP like.
 DR InterPro; IPR000873; AMP-bind.
 DR InterPro; IPR006163; Phosphatase.
 DR Pfam; PF00501; AMP-binding; 1.
 DR Pfam; PF00550; PP-binding; 1.
 DR PRINTS; PR00154; AMPBINDING.

```

DR PROSITE; PSS0075; ACP_DOMAIN; 1.
KW Ligase.
SQ SEQUENCE 705 AA; 75197 MW; 600F2D0EABFDF1DC CRC64;

Query Match 60.3%; Score 44; DB 2; Length 705;
Best Local Similarity 70.0%; Pred. No. 37;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 YRCRGDDSKV 11
   ||||| :|||
Db 383 YRCRADDTTEV 392

RESULT 11
Q7TXK7 PRELIMINARY; PRT; 705 AA.
AC Q7TXK7;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 26, Last annotation update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PROBABLE FATTY-ACID-CoA LIGASE FADD22 (FATTY-ACID-CoA SYNTHETASE)
DE (FATTY-ACID-CoA SYNTHASE) (EC 6.2.1.-.-)
GN Names-fad22; OrderedLocusNames=Mb2972c;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OC NCBI_TaxID=1765;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eiglmeyer K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrall B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR EMBL; BX248344; CAD96659.1; -.
DR GO; GO:0048037; F:cofactor binding; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR009081; ACP like.
DR InterPro; IPR000873; AMP-bind.
DR InterPro; IPR006163; Phspanteth_bind.
DR Pfam; PF00501; AMP-binding; 1.
DR Pfam; PF00550; PP-binding; 1.
DR PROSITE; PSS0075; ACP_DOMAIN; 1.
KW Complete proteome; Ligase.
SQ SEQUENCE 705 AA; 75197 MW; 600F2D0EABFDF1DC CRC64;

Query Match 60.3%; Score 44; DB 2; Length 705;
Best Local Similarity 70.0%; Pred. No. 37;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 YRCRGDDSKV 11
   ||||| :|||
Db 383 YRCRADDTTEV 392

RESULT 12
Q6AX87 PRELIMINARY; PRT; 307 AA.
AC Q6AX87;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE MGC81990 protein.
GN Name=MGC81990;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OC NCBI_TaxID=8355;

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalios D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX Klein S., Gerhard D.S.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC079709; AAH79709.1; -.
DR InterPro; IPR008934; ACPase_VanPerase.
DR InterPro; IPR000326; Pesterase_PA_PTP.
DR Pfam; PF01569; PAP2; 1.
DR SMART; SM00014; acidPPC; 1.
SQ SEQUENCE 307 AA; 34417 MW; 23P956E7B8FFED0D CRC64;

Query Match 58.9%; Score 43; DB 2; Length 307;
Best Local Similarity 61.5%; Pred. No. 23;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 NYRCRGDDSKVQ 13
   ||||| :|||
Db 179 NYRCRGPFNKWE 191

RESULT 13
Q7QP19 PRELIMINARY; PRT; 73 AA.
AC Q7QP19;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE GLP 83 1716 1495.
OS Giardia lamblia ATCC 50803.
OC Eukaryota; Diplomonadida; Hexamitidae; Giardiinae; Giardia.
OC NCBI_TaxID=184922;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WB C6;
RA Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D.,
RA Olsen G.J., Sogin M.L.;
RT "Draft sequence of the Giardia lamblia genome.";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an

```

```

CC      EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC      preliminary data.
DR      EMBL; AACB01000181; EAA36741.1; -.
SQ      SEQUENCE   73 AA;  8125 MW;  4F866C9AEF254DDE CRC64;

Query Match          57.5%; Score 42; DB 2; Length 73;
Best Local Similarity 75.0%; Pred. No. 7.3; Indels 0; Gaps 0;
Matches 6; Conservative 2; Mismatches 0;

Qy      1 NYRCRGDD 8
Db      27 DYRCCKGDD 34

RESULT 14
ID      Q86KQ7      PRELIMINARY;      PRT;      265 AA.
AC      Q86KQ7;
DT      01-JUN-2003 (TrEMBLrel. 24, Created)
DT      01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Hypothetical protein.
OS      Dictyostelium discoideum (Slime mold).
OC      Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX      NCBI_TaxID=44689;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=AX4;
RX      MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
RA      Gloeckner G., Bichinger L., Szafranski K., Pachebat J., Dear P.,
RA      Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA      Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT      "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
RL      Nature 418:79-85(2002).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=AX4;
RA      Baumgart C.;
RL      Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AC115584; AA050949.1; -.
DR      InterPro; IPR008997; RicinB_like.
KW      Hypothetical protein.
SQ      SEQUENCE   265 AA;  30503 MW;  F1E238D4A048F6D2 CRC64;

Query Match          57.5%; Score 42; DB 2; Length 265;
Best Local Similarity 87.5%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 YRCRGDD 9
Db      45 YRCRNDSD 52

RESULT 15
Q9A731
ID      Q9A731      PRELIMINARY;      PRT;      513 AA.
AC      Q9A731;
DT      01-JUN-2001 (TrEMBLrel. 17, Created)
DT      01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Anthranilate synthase component I.
GN      OrderedLocNames=CC1895;
OS      Caulobacter crescentus.
OC      Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
OC      Caulobacteraceae; Caulobacter.
OX      NCBI_TaxID=155892;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=ATCC 19089 / CB15;
RX      MEDLINE=21173698; PubMed=11259647; DOI=10.1073/pnas.061029298;
RA      Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA      Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA      Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,

```

```

RA      DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA      Kolonay J.F., Smit J., Craven M.B., Khouri H.M., Shetty J.,
RA      Berry K.J., Utterback T.R., Tran K., Wolf A.M., Vamathevan J.J.,
RA      Ermolaeva M.D., White O., Salzberg S.L., Venter J.C., Shapiro L.,
RA      Fraser C.M.;
RT      "Complete genome sequence of Caulobacter crescentus.";
RL      Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR      EMBL; AS005863; AAK23870.1; -.
DR      PIR; B87484; B87484.
DR      HSSP; P05041; 1KOG.
DR      TIGR; CC1895; -.
DR      GO; GO:0004049; F:anthranilate synthase activity; IEA.
DR      GO; GO:0009058; P:biogenesis; IEA.
DR      GO; GO:0000162; P:tryptophan biosynthesis; IEA.
DR      InterPro; IPR005256; Anth synth.
DR      InterPro; IPR005801; Anth synth chor.
DR      InterPro; IPR006805; Anth synth I.N.
DR      Pfam; PF04715; Anth synt_I_N_1.
DR      PRINTS; PR00095; Anth synth chor.
DR      TIGRFAMs; TIGR00564; trpE_most; 1.
KW      Complete proteome.
SQ      SEQUENCE   513 AA;  55794 MW;  A02E026EE08AE30C CRC64;

Query Match          57.5%; Score 42; DB 2; Length 513;
Best Local Similarity 50.0%; Pred. No. 61;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy      2 YRCRGDDSKVQ 13
Db      73 WRCRGDQAEIAE 84

RESULT 16
Q8WQJ3
ID      Q8WQJ3      PRELIMINARY;      PRT;      845 AA.
AC      Q8WQJ3;
DT      01-MAR-2002 (TrEMBLrel. 20, Created)
DT      01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT      01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE      SRCR LCCL adhesive-like protein (Fragment).
GN      Name=slap;
OS      Plasmodium chabaudi.
OC      Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX      NCBI_TaxID=5825;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=21982590; PubMed=11985859; DOI=10.1016/S0166-6851(02)00016-6;
RA      Delrieu I., Waller C.C., Mota M.M., Grainger M., Langhorne J.,
RA      Holder A.A.;
RT      "PSLAP, a protein with multiple adhesive motifs, is expressed in
RT      Plasmodium falciparum gametocytes.";
RL      Mol. Biochem. Parasitol. 121:11-20(2002).
DR      EMBL; AY072023; AAL58521.1; -.
DR      GO; GO:0016020; C:membrane; IEA.
DR      GO; GO:0005044; F:scavenger receptor activity; IEA.
DR      Pfam; PF03815; LCCL; 3.
DR      Pfam; PF0530; SRCR; 2.
DR      SMART; SM00202; SR; 2.
DR      PROSITE; PS50820; LCCL; 2.
DR      PROSITE; PS50287; SRCR_2; 2.
FT      NON TER 1
FT      NON TER 1
SQ      SEQUENCE   845 AA;  95032 MW;  06A389B4CC294477 CRC64;

Query Match          57.5%; Score 42; DB 2; Length 845;
Best Local Similarity 41.7%; Pred. No. 16+02;
Matches 5; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

Qy      1 NYRCRGDDSKVQ 12
Db      340 NFRCKGDEANLK 351

```

```
RESULT 17
Q8WT63      PRELIMINARY;      PRT; 1304 AA.
AC Q8WT63;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Multidomain scavenger receptor protein PSBR precursor.
OS Plasmodium berghei.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5821;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22242506; PubMed=12354219;
RA Claudiano C., Dessens J.T., Trueman H.E., Arai M., Mendoza J.,
RA Butcher G.A., Crompton T., Sindén R.E.;
RA "A malaria scavenger receptor-like protein essential for parasite
RT development.";
RL Mol. Microbiol. 45:1473-1484 (2002).
CC -!- SIMILARITY: Contains 1 PLAT domain.
DR EMBL: AV034780; AAK64185.1; -.
DR HSSP; O43405; 1JBI.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0005044; F:scavenger receptor activity; IEA.
DR Pfam; PF03815; LCC1; 4.
DR Pfam; PF01477; FLAT; 1.
DR Pfam; PF00530; SRCR; 2.
DR SMART; SM00202; SR; 2.
DR PROSITE; PS50820; LCC1; 3.
DR PROSITE; PS50095; PLAT; 1.
DR PROSITE; PS50287; SRCR_2; 2.
KW Receptor; Signal.
FT SIGNAL 1 22 Potential.
SQ SEQUENCE 1304 AA; 148247 MW; F936CC94E7B19F1C CRC64;

Query Match 57.5%; Score 42; DB 2; Length 1304;
Best Local Similarity 41.7%; Pred. No. 1.7e+02;
Matches 5; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDSDSKVQ 12
:|:|:|:|:|:
DB 614 NFRCKGDEANLK 625

RESULT 18
Q8POM2      PRELIMINARY;      PRT; 37 AA.
AC Q8POM2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical phage protein.
GN OrderedLocustNames=spym18.1296;
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=186103;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MGAS8232;
RA MEDLINE=21927593; PubMed=11917108; DOI=10.1073/pnas.062526099;
RA Smoot J.C., Barbian K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylvia G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RA "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks.";
RA Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673 (2002).
DR EMBL; AB010051; AAL97902.1; -.

KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 37 AA; 4398 MW; E7E884FEAA642342 CRC64;

Query Match 56.2%; Score 41; DB 2; Length 37;
Best Local Similarity 61.5%; Pred. No. 5.3;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 NYRCRGDSDSKVQE 13
|||:|:|:|
DB 14 NYHTNGDDSKYLE 26

RESULT 19
Q69XV4      PRELIMINARY;      PRT; 226 AA.
AC Q69XV4;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE DNA-binding protein-like.
GN Name=P0040H10.36-2;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC
RT clone:P0040H10.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003539; BAD35373.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
KW DNA-binding.
SQ SEQUENCE 226 AA; 24068 MW; 3797E1C36CC9A244 CRC64;

Query Match 56.2%; Score 41; DB 2; Length 226;
Best Local Similarity 53.8%; Pred. No. 38;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 NYRCRGDSDSKVQE 13
:|:|:|:|:|:
DB 153 DYSCGGDDSGTGE 165

RESULT 20
Q8TBF1      PRELIMINARY;      PRT; 428 AA.
AC Q8TBF1;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE RasGEF domain family, member 1A.
GN Name=RASGEF1A;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Brain;
RA MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
```



```

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Heiton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skaleka U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Maria M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RL Strauberg R.;
RL Submitted (Feb-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC022548; AAt2548.1; -.
DR GO; GO:0005085; F:guanyl-nucleotide exchange factor activity; IEA.
DR GO; GO:0007264; P:small GTPase mediated signal transduction; IEA.
DR InterPro; IPR000651; RasGEF_N.
DR InterPro; IPR001895; RasGRF_CDC25.
DR InterPro; IPR008937; Ras_GEF.
DR Pfam; PF00617; RasGEF; 1.
DR SMART; SM00147; RasGEF; 1.
DR PROSITE; PS50009; RASGEF_CAT; 1.
DR PROSITE; PS50212; RASGEF_NTER; 1.
SQ SEQUENCE 428 AA; 49394 MW; 2522DBF4A3D2F560 CRC64;

Query Match 56.2%; Score 41; DB 2; Length 428;
Best Local Similarity 70.0%; Pred. No. 76;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDSDK 10
Db 197 NHRCRGDLTK 206

RESULT 21
ID Q69XV5 PRELIMINARY; PRT; 461 AA.
AC Q69XV5;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE DNA-binding protein-like.
GN Names=P0040H10.36-1;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC
RT clone:P0040H10.36-1;
RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003539; BAD35372.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR001605; ARID.
DR InterPro; IPR002068; Hsp20.
DR InterPro; IPR008978; HSP20_chap.
DR Pfam; PF01388; ARID; 1.
DR SMART; SM00501; BRIGHT; 1.
DR PROSITE; PS01031; HSP20; 1.
DR DNA-binding.
SQ SEQUENCE 461 AA; 50459 MW; F6FD7A3F1B13290B CRC64;

Query Match 56.2%; Score 41; DB 2; Length 461;
Best Local Similarity 53.8%; Pred. No. 82;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 NYRCRGDSDKVQE 13

```

```

Db 153 DYSCGDDSGTGE 165
:| | | | |
:| | | | |

RESULT 22
ID Q8N9B8 PRELIMINARY; PRT; 489 AA.
AC Q8N9B8;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein FLJ37817.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Substantia nigra;
RX PubMed=14702039; DOI=10.1038/ngl285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Katsuma N., Sato K., Tanikawa M., Yamazaki M.,
RA Abe K., Kimihara K., Katsuta N., Sato K., Tanikawa M., Fujimori K.,
RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Chiba Y., Ishida S.,
RA Tanai H., Kimata M., Watanabe M., Hiraoa S., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yoshida M., Hotuta T., Kusano J.,
RA Kanehori K., Takahashi-Fujii A., Hara R., Takeuchi K., Arita M., Imose N.,
RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiya H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi R.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RN Nat. Genet. 36:40-45(2004).
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC
RT clone:P0040H10.36-1;
RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK095136; BAC04491.1; -.
DR GO; GO:0005085; F:guanyl-nucleotide exchange factor activity; IEA.
DR GO; GO:0007264; P:small GTPase mediated signal transduction; IEA.
DR InterPro; IPR000651; RasGEF_N.
DR InterPro; IPR001895; RasGRF_CDC25.
DR InterPro; IPR008937; Ras_GEF.
DR Pfam; PF00617; RasGEF; 1.
DR SMART; SM00147; RasGEF; 1.
DR SMART; SM00229; RasGEF_N.
DR PROSITE; PS50009; RASGEF_CAT; 1.
DR PROSITE; PS50212; RASGEF_NTER; 1.
SQ SEQUENCE 489 AA; 55532 MW; 85D5986ACD2C7C76 CRC64;

Query Match 56.2%; Score 41; DB 2; Length 489;
Best Local Similarity 70.0%; Pred. No. 87;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDSDK 10
Db 258 NHRCRGDLTK 267
:| | | | |
:| | | | |

```

RESULT 23

```
Q96180
ID Q96180 PRELIMINARY; PRT; 505 AA.
AC Q96180;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PRSS12 protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zebberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimlock J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek A., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC007761; AA07761.1; -.
DR HSP; Q08380; IBY2.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005507; F:copper ion binding; IEA.
DR GO; GO:0004720; F:protein-lysine 6-oxidase activity; IEA.
DR GO; GO:0005044; F:scavenger receptor activity; IEA.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001695; Lysyl oxidase.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00530; SRCR; 3.
DR PRINTS; PRO00018; KRINGLE.
DR PRINTS; PRO0258; SPERACTRCPPR.
DR ProDom; PD000395; Kringle; 1.
DR ProDom; PD013887; Lysyl_oxidase; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00202; SR; 3.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00420; SRCR_1; 2.
DR PROSITE; PS00287; SRCR_2; 3.
KW Kringle.
SQ SEQUENCE 505 AA; 55658 MW; 7FF863A6246226BD CRC64;
```

Query Match 56.2%; Score 41; DB 2; Length 505;

Best Local Similarity 54.5%; Pred. No. 90;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 NVRCRGDSKV 11

Db 236 NVRCRGDEINI 246

RESULT 24

```
NETR_HUMAN
ID P56730; Q9UP16; PRT; 875 AA.
AC P56730; Q9UP16;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Neurotysin precursor (EC 3.4.21.-) (Motopsin) (Leydin).
GN Name=PRSS12;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98201705; PubMed=9540828; DOI=10.1016/S0167-4781(97)00205-4;
RA Proba K., Gschwend T.P., Sonderegger P.;
RT "Cloning and sequencing of the cDNA encoding human neurotysin.";
RL Biochim. Biophys. Acta 1396:143-147 (1998).
RN [2]
RP SEQUENCE OF 615-875 FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=99203523; PubMed=10103056;
RA Poorafshar M., Hellman L.;
RT "Cloning and structural analysis of leydin, a novel human serine
RT protease expressed by the Leydig cells of the testis.";
RL Eur. J. Biochem. 261:244-250 (1999).
RN [3]
RP DISEASE.
RX PubMed=12459588; DOI=10.1126/science.1076521;
RA Molinari F., Rio M., Meskenaitė V., Encha-Razavi F., Auge J., Bacq D.,
RA Briault S., Vekemans M., Munnich A., Attie-Bitach T., Sonderegger P.,
RA Collea L.;
RT "Truncating neurotysin mutation in autosomal recessive nonsyndromic
RT mental retardation.";
RL Science 298:1779-1781 (2002).
CC -!- FUNCTION: Plays a role in neuronal plasticity and the proteolytic
CC action may subserve structural reorganizations associated with
CC learning and memory operations (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DISEASE SPECIFICITY: Brian and Leydig cells of the testis.
CC -!- DISEASE: Defects in PRSS12 are a cause of autosomal recessive
CC nonsyndromic mental retardation [MIM:249500]. Mental retardation
CC is a mental disorder characterized by significantly subaverage
CC general intellectual functioning associated with impairments in
CC adaptive behavior and manifested during the developmental
CC period.
CC -!- SIMILARITY: Belongs to the peptidase S1 family.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 4 SRCR domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ001531; CAA04816.1; -.
DR EMBL; AF077298; AAD25919.1; -.
DR HSP; P00760; IEZX.
DR MEROPS; S01.237; -.
DR Genew; HGNC:9477; PRSS12.
DR MIM; 249500; -.
DR MIM; 606709; -.
DR GO; GO:0008236; F:serine-type peptidase activity; TAS.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
```

DR InterPro: IPR001190; Srcr_receptor.
 DR Pfam: PF00051; Kringle; 1.
 DR PFam: PF00530; SRCR; 4.
 DR PFam: PF00089; Trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR00258; SPERACTRCPTR.
 DR ProDom: PD000395; Kringle; 1.
 DR ProDom: PD013887; Lysyl_oxidase; 2.
 DR SMART: SM00130; KR; 1.
 DR SMART: SM00202; SR; 4.
 DR SMART: SM00020; Tryp SPC; 1.
 DR PROSITE: PS00021; KRINGLE 1; 1.
 DR PROSITE: PS00720; KRINGLE_2; 1.
 DR PROSITE: PS00420; SRCR_1; 3.
 DR PROSITE: PS00420; SRCR_2; 4.
 DR PROSITE: PS00287; SRCR_2; 4.
 DR PROSITE: PS02040; TRYPSIN DOM; 1.
 DR PROSITE: PS00134; TRYPSIN HM; 1.
 DR PROSITE: PS00135; TRYPSIN SR; 1.
 KW Glycoprotein; Hydrolase; Kringle; Repeat; Serine protease; Signal.
 FT SIGNAL 1 20 Potential.
 FT CHAIN 21 875 Neurotrypsin.
 FT DOMAIN 23 92 Pro-rich.
 FT DOMAIN 93 165 Kringle.
 FT DOMAIN 170 271 SRCR 1.
 FT DOMAIN 280 381 SRCR 2.
 FT DOMAIN 387 487 SRCR 3.
 FT DOMAIN 500 601 SRCR 4.
 FT DOMAIN 619 875 Serine protease.
 FT DOMAIN 619 630 Zymogen activation region.
 FT SITE 630 631 Reactive bond (Potential).
 FT ACT_SITE 676 676 Charge relay system.
 FT ACT_SITE 726 726 Charge relay system.
 FT ACT_SITE 825 825 Charge relay system.
 FT DISULFID 619 750 Potential.
 FT CARBOHYD 26 26 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 683 683 A-linked (GlcNAc...) (Potential).
 FT CONFLICT 663 663 A -> V (in Ref. 2).
 FT CONFLICT 701 701 E -> V (in Ref. 2).
 FT CONFLICT 839 841 VVY -> AAL (in Ref. 2).
 SQ SEQUENCE 875 AA; 97011 MW; B66EC946DC208DC8 CRC64;

Query Match 56.2%; Score 41; DB 1; Length 875;
 Best Local Similarity 54.5%; Pred. No. 1.6e+02;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 NVRCRGDSKV 11
 |||||:
 Db 236 NVRCRGDENI 246

RESULT 25
 O97322 PRELIMINARY; PRT; 1334 AA.
 AC O97322; (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 25, Last annotation update)
 DE Hypothetical protein MAL3P7.32.
 GN Name=MAL3P7.32; Synonyms=PF01010w;
 OS Plasmodium falciparum (isolate 3D7).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=36329;
 RN SEQUENCE FROM N.A.
 RX MEDLINE=99376085; PubMed=10448855; DOI=10.1038/22964;
 RA Bowman S., Lawson D., Basham D., Brown D., Chillingworth T.,
 RA Churcher C.M., Craig A., Davies R.M., Devlin K., Felwell T.,
 RA Gentles S., Gwilliam R., Hamlin N., Harris D., Holroyd S., Hornsby T.,
 RA Horrocks P., Jagels K., Jaseal B., Kyes S., McLean J., Moule S.,
 RA Mungall K., Murphy L., Oliver K., Quail M.A., Rajandream M.-A.,
 RA Rutter S., Skelton J., Squares R., Squares S., Sulston J.E.,
 RA Whitehead S., Woodward J.R., Newbold C., Barrell B.G.;
 RT "The complete nucleotide sequence of chromosome 3 of Plasmodium

RT falciparum.";
 RL Nature 400:532-538 (1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22255708; PubMed=12368867; DOI=10.1038/nature01095;
 RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
 RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
 RA Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,
 RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
 RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
 RA Feltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
 RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
 RA Humphray S., Jagels K., James K.D., Johnson D., Keshornou A.,
 RA Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,
 RA Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,
 RA Oliver K., Ormond D., Price C., Quail M.A., Rabinowitsch E.,
 RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
 RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
 RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
 RA Sulston J.E., Craig A., Newbold C., Barrell B.G.;
 RL "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13.";
 RL Nature 419:527-531 (2002).
 DR EMBL: AL034559; CAB39073.2; -.
 KW Hypothetical protein.
 SQ SEQUENCE 1334 AA; 160847 MW; E3577E84C7E0C8E5 CRC64;

Query Match 56.2%; Score 41; DB 2; Length 1334;
 Best Local Similarity 60.0%; Pred. No. 2.6e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NVRCRGDSK 10
 |||||:
 Db 808 NYNCKDDNK 817

RESULT 26
 Q9RTC3 PRELIMINARY; PRT; 2736 AA.
 AC Q9RTC3; (TREMBLrel. 26, Created)
 DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Hypothetical protein.
 GN Name=PY00071;
 OS Plasmodium yoelii yoelii.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=73239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=17XNL;
 RX PubMed=12368865; DOI=10.1038/nature01099;
 RA Carlton J.M., Anguoli S.V., Suh B.B., Kooij T.W., Perte M.,
 RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
 RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
 RA Shallom S.J., van Aken S.E., Riedmiller S.B., Feidlyum T.V.,
 RA Cho J.K., Quackenbush J., Sedegah M., Shoabi A., Cummings L.M.,
 RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
 RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
 RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
 RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
 RA Carucci D.J.;
 RT "Genome sequence and comparative analysis of the model rodent malaria
 parasite Plasmodium yoelii yoelii.";
 RL Nature 419:512-519 (2002).
 CC -!- CAUTION: The sequence shown here is derived from an
 EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 preliminary data.
 CC EMBL: AAL010000.9; EAA17894.1; -.
 DR InterPro: IPR000357; HEAT.
 DR Pfam: PF02985; HEAT; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 2736 AA; 323770 MW; C219F7662538AA00 CRC64;

Query Match 56.2%; Score 41; DB 2; Length 2736;
Best Local Similarity 46.2%; Pred. No. 5.6e+02;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 NYRCRGDDSKVQE 13
DB 2536 NYACSTDQKIKD 2548
||| ||| ||| :| :| :| :|

RESULT 27
Q9KCP2 SEQUENCE FROM N.A.
ID Q9KCP2 PRELIMINARY; PRT; 150 AA.
AC Q9KCP2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Ferric uptake regulation protein.
GN OrderedLocusNames=BHI527;
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125;
RX MEDLINE=20512582; PubMed=11058132; DOI=10.1093/nar/28.21.4317;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL; AP001512; BAB05246.1; -.
DR PIR; G83840; G83840.
DR HSP; Q03456; LMZB.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR002481; FUR.
DR InterPro; IPR009058; wing_hlx_DNA_bnd.
DR Pfam; PF01475; FUR; 1.
DR ProDom; PD002003; FUR; 1.
KW Complete proteome.
SQ SEQUENCE 150 AA; 17447 MW; CE4D3AC2A8B64275 CRC64;

Query Match 54.8%; Score 40; DB 2; Length 150;
Best Local Similarity 66.7%; Pred. No. 37;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDDSK 10
DB 141 HRCQGDESK 149
:| :| :| :|

RESULT 28
Q20683 SEQUENCE FROM N.A.
ID Q20683 PRELIMINARY; PRT; 179 AA.
AC Q20683;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein F52H3.5.
GN ORFNames=F52H3.5;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology.";
RL Science 282:2012-2018(1998).

RN SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Gardner A.B.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z66512; CAA91325.1; -.
DR PIR; T22521; T22521.
DR WormBase; WBGene00009947; F52H3.5.
DR WormPep; F52H3.5; CE03401.
DR InterPro; IPR001440; TPR.
DR InterPro; IPR008941; TPR-like.
DR Pfam; PF00515; TPR_1; 1.
DR SMART; SM00028; TPR; 3.
DR PROSITE; PS00005; TPR; 1.
DR PROSITE; PS50293; TPR_REGION; 1.
KW Hypothetical protein; Repeat; TPR repeat.
SQ SEQUENCE 179 AA; 19648 MW; 6FF899837C63A18F CRC64;

Query Match 54.8%; Score 40; DB 2; Length 179;
Best Local Similarity 63.6%; Pred. No. 45;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 YRCRGDDSKVQ 12
DB 125 YRLRGDDDKAR 135
||| ||| :| :| :|

RESULT 29
Q95ZL6 SEQUENCE FROM N.A.
ID Q95ZL6 PRELIMINARY; PRT; 244 AA.
AC Q95ZL6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein M01E11.4.
GN Name=M01E11.4; ORFNames=M01E11.4;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG WormBase Consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Pauley A., Gattung S.;
RT "The sequence of C. elegans cosmid M01E11.";
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RG WormBase Consortium;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; U80450; AAK77613.1; -.
DR WormBase; WBGene00019713; M01E11.4.
DR WormPep; M01E11.4a; CE28610.
KW Hypothetical protein.

```
SQ SEQUENCE 244 AA; 27837 MW; 9AA335B0442B84C4 CRC64;
Query Match 54.8%; Score 40; DB 2; Length 244;
Best Local Similarity 70.0%; Pred. No. 62;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 CRGDDSKVOE 13
   ||| ||| |||
Db 42 CRGLDSQIQE 51

RESULT 30
ID Q95ZL7 PRELIMINARY; PRT; 252 AA.
AC Q95ZL7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein M01E11.4.
GN Names=M01E11.4; ORNames=M01E11.4;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=Bristol N2;
RC MEDLINE=99069613; PubMed=9851916;
RG WormBase Consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium. ";
RL Science 282:2012-2018 (1998).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=Bristol N2.
RC STRAIN=Bristol N2;
RA Pauley A., Gattung S.;
RT "The sequence of C. elegans cosmid M01E11. ";
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.
RP STRAIN=Bristol N2;
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
[4]
RN SEQUENCE FROM N.A.
RP STRAIN=Bristol N2;
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
[5]
RN SEQUENCE FROM N.A.
RP STRAIN=Bristol N2;
RC WormBase Consortium;
RG WormBase Consortium;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; U80450; AAK77614.1; -.
DR WormBase; WBGene00019713; M01E11.4.
DR WormPep; M01E11.4b; CE28611.
KW Hypothetical protein.
SQ SEQUENCE 252 AA; 28738 MW; FE8556D34587F3E8 CRC64;

Query Match 54.8%; Score 40; DB 2; Length 252;
Best Local Similarity 70.0%; Pred. No. 65;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 CRGDDSKVOE 13
   ||| ||| |||
Db 42 CRGLDSQIQE 51

RESULT 31
Q7R8M3 PRELIMINARY; PRT; 269 AA.
ID Q7R8M3
AC Q7R8M3;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
```

```
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
GN Name=Yir3 protein.
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=73239;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=17XNL;
RC PubMed=12368865; DOI=10.1038/nature01099;
RA Carlton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Perteau M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoabi A., Cummings L.M.,
RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria
RT parasite Plasmodium yoelii yoelii. ";
RL Nature 419:512-519 (2002).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABL01002583; EAA19571.1; -.
DR InterPro; IPR006477; Yir_bir_cir.
DR Pfam; PF06022; Cir_Bir_Yir_I.
DR TIGRFAMs; TIGR01590; Yir-bir-cir_pla; 1.
SQ SEQUENCE 269 AA; 31732 MW; A95E52E05D6119C CRC64;

Query Match 54.8%; Score 40; DB 2; Length 269;
Best Local Similarity 54.5%; Pred. No. 69;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 NYRCRGDDSKV 11
   ||| ||| |||
Db 38 NYKCEGLDKI 48

RESULT 32
LPPI_FAT
ID LPPI_FAT STANDARD; PRT; 282 AA.
AC O08564; Q8K594;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Lipid phosphate phosphohydrolase 1 (EC 3.1.3.4) (Phosphatidic acid
DE phosphatase 2a) (Phosphatidate phosphohydrolase type 2a) (PAP2a) (PAP-
DE 2a) (PAP2-alpha).
GN Name=Ppap2a; Synonyms=Lpp1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
RN SEQUENCE FROM N.A. (ISOFORM 1), AND CHARACTERIZATION.
RP TISSUE=Liver;
RC PubMed=10359651; DOI=10.1042/0264-6021:3400677;
RA Jasinska R., Zhang Q.-X., Pilquil C., Singh I., Xu J., Dewald J.,
RA Dillon D.A., Berthiaume L.G., Carman G.M., Waggoner D.W.,
RA Brindley D.N.;
RT "Lipid phosphate phosphohydrolase-1 degrades exogenous glycerolipid
RT and sphingolipid phosphate esters. ";
RL Biochem. J. 340:677-686 (1999).
[2]
RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RP STRAIN=Sprague-Dawley; TISSUE=Lung;
RC MEDLINE=21559599; PubMed=11704545;
RA Nanjundan M., Possmayer F.;
RT "Molecular cloning and expression of pulmonary lipid phosphate
```

phosphohydrolases.";
Am. J. Physiol. 281:L1484-L1493 (2001).
[3]
CHARACTERIZATION
PubMed=8663293; DOI=10.1074/jbc.271.28.16506;
Waggoner D.W., Gomez-Munoz A., Dewald J., Brindley D.N.;
"Phosphatidate phosphohydrolase catalyzes the hydrolysis of ceramide
1-phosphate, lysophosphatidate, and sphingosine 1-phosphate.";
J. Biol. Chem. 271:16506-16509(1996).
-/- FUNCTION: Broad-specificity phosphohydrolase that dephosphorylates
exogenous bioactive glycerolipids and sphingolipids. Catalyzes the
conversion of phosphatidic acid (PA) to diacylglycerol (DG). In
addition it hydrolyzes lysophosphatidic acid (LPA), diacyl
glycerol pyrophosphate (DGPP), ceramide-1-phosphate (C-1-P) and
sphingosine-1-phosphate (S-1-P). The relative catalytic efficiency
is LPA > PA > C-1-P > S-1-P.
-/- CATALYTIC ACTIVITY: A 3-sn-phosphatidate + H(2)O = a 1,2-diacyl-
sn-glycerol + phosphate.
-/- SUBUNIT: Homodimer (By similarity).
-/- SUBCELLULAR LOCATION: Integral membrane protein. Plasma membrane
(By similarity).
-/- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
Comment=Additional isoforms seem to exist;
Name=1;
IsoId=O08564-1; Sequence=Displayed;
Name=2; Synonyms=LpP1a;
IsoId=O08564-2; Sequence=VSP_009653;
-/- SIMILARITY: Belongs to the PA-phosphatase related phosphoesterase
family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; U90556; AAB50246.1; --
EMBL; AF503609; AM28631.1; --
GO; GO:0005624; C=membrane fraction; ISS.
GO; GO:0008195; F:phosphatidate phosphatase activity; ISS.
GO; GO:0030521; P:androgen receptor signaling pathway; ISS.
GO; GO:0008285; P:negative regulation of cell proliferation; ISS.
GO; GO:0046839; P:phospholipid dephosphorylation; ISS.
GO; GO:0007205; P:protein kinase C activation; ISS.
GO; GO:0019216; P:regulation of lipid metabolism; ISS.
InterPro; IPR008934; ACPase_VanPase.
Pfam; PF01569; PAP2; 1.
SMART; SM00014; acidPpc; 1.
Alternative splicing; Glycoprotein; Hydrolase; Transmembrane.
DOMAIN 1 6 Cytoplasmic (Potential).
TRANSMEM 7 27 Potential.
DOMAIN 28 53 Extracellular (Potential).
TRANSMEM 54 74 Potential.
DOMAIN 75 88 Cytoplasmic (Potential).
TRANSMEM 89 109 Potential.
DOMAIN 110 164 Extracellular (Potential).
TRANSMEM 165 185 Potential.
DOMAIN 186 194 Cytoplasmic (Potential).
TRANSMEM 195 215 Potential.
DOMAIN 216 229 Extracellular (Potential).
TRANSMEM 230 250 Potential.
DOMAIN 251 282 Cytoplasmic (Potential).
CARBOHYD 142 142 N-linked (GlcNAc...) (Potential).
VARSPIC 21 70
IPVFCIIIV -> SNPMVNLGQIYPFORGFCSDNSVKY
PYHDSVTVTSVLVLGLGPIPIFS (in isoform 2).
/FTId=VSP_009653.
SEQUENCE 282 AA; 31996 MW; A4ED3DEB33FD7943 CRC64;

Query Match 54.8%; Score 40; DB 1; Length 282;
Best Local Similarity 46.2%; Pred. No. 73;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
QY 1 NYRCRGDDSKVQE 13
Db 150 NFVQCQNEQKVRE 162
|:|:|:|:|:
RESULT 33
Q6P766 PRELIMINARY; PRT; 282 AA.
AC Q6P766;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max A.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ussin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Strausberg R.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC061815; AAH61815.1; --
DR InterPro; IPR008934; ACPase_VanPase.
DR InterPro; IPR000326; Pesterase_PA_PTP.
DR Pfam; PF01569; PAP2; 1.
DR SMART; SM00014; acidPpc; 1.
KW Hypothetical protein.
SQ SEQUENCE 282 AA; 32000 MW; 10EFD1C5F3FCB8E9 CRC64;
Query Match 54.8%; Score 40; DB 2; Length 282;
Best Local Similarity 46.2%; Pred. No. 73;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
QY 1 NYRCRGDDSKVQE 13
Db 150 NFVQCQNEQKVRE 162
|:|:|:|:|:
RESULT 34
Q9DW23 PRELIMINARY; PRT; 288 AA.
ID Q9DW23
AC Q9DW23;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)

DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE P171.
GN Name=171;
OS Rat cytomegalovirus (strain Maastricht).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Muromegalovirus.
OX NCBI_TaxID=79700;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Maastricht;
RX MEDLINE=96357047; PubMed=8764031;
RA Vink C., Beuken E., Bruggeman C.A.;
RT "Structure of the rat cytomegalovirus genome termini.";
RL J. Virol. 70:5221-5229(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Maastricht;
RX MEDLINE=20366325; PubMed=10906222;
RA Vink C., Beuken E., Bruggeman C.A.;
RT "Complete DNA sequence of the rat cytomegalovirus genome.";
RL J. Virol. 74:7656-7665(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Maastricht;
RX MEDLINE=20473137; PubMed=11018281; DOI=10.1016/S0168-1702(00)00208-2;
RA Gruijthuisen Y.K., Beuken E., Bruggeman C.A.;
RT "Rat cytomegalovirus R89 is a highly conserved gene which expresses a
RT spliced transcript.";
RL Virus Res. 69:119-130(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Maastricht;
RA Vink C., Beuken E., Bruggeman C.A.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF232689; AAF99267.1; - B29EDAF6083D698B CRC64;
SQ SEQUENCE 288 AA; 32646 MW; B29EDAF6083D698B CRC64;
Query Match 54.8%; Score 40; DB 2; Length 288;
Best Local Similarity 60.0%; Pred. No. 75;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 3 RCRGDSKVQ 12
DB 138 RCQGEDEKVE 147
|||||:
RESULT 35
HEM3_XANCP STANDARD; PRT; 304 AA.
AC QRP536;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Porphobilinogen deaminase (EC 2.5.1.61) (PBG) (Hydroxymethylbilane
DE synthase) (HMB) (pre-uroporphyrinogen synthase).
GN Name=hmc; OrderedLocustNames=XCC3511;
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=22022145; PubMed=12024217; DOI=10.1038/417459a;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavaro F., Cardoso J., Chamberg F., Chapina L.P.,
RA Ciccarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,

RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Teai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities";
RL Nature 417:459-463(2002).
CC -!- FUNCTION: Tetrapolymerization of the monopyrrole PBG into the
CC hydroxymethylbilane preuroporphyrinogen in several discrete steps.
CC -!- CATALYTIC ACTIVITY: 4 porphobilinogen + H2O =
CC hydroxymethylbilane + 4 NH(3).
CC -!- COFACTOR: Covalently binds a dipyrromethane cofactor to which the
CC porphobilinogen subunits are added (By similarity).
CC -!- PATHWAY: Porphyrin biosynthesis by the C5 pathway; fourth step.
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SIMILARITY: Belongs to the HMBs family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AB012470; AAM42781.1; -
DR HSSP: P06983; IYPN.
DR HAMAP: MF_00260; -; 1.
DR InterPro: IPR000860; Porphobil_deam.
DR Pfam: PF01379; Porphobil_deam; 1.
DR Pfam: PF03900; Porphobil_deamC; 1.
DR PRINTS: PR00151; PORPHBDMNASE.
DR ProDom: PD002745; Porphobil_deam; 1.
DR TIGRPFAMs: TIGR00212; hmc; 1.
DR PROSITE: PS00533; PORPHOBILINOGEN_DEAM; 1.
DR Complete proteome: Porphyrin biosynthesis; Transferase.
KW Complete proteome: Porphyrin biosynthesis; Transferase.
FT BINDING 240 240 Pyrromethane cofactor (By similarity).
SQ SEQUENCE 304 AA; 32460 MW; F77A49FB1932FF02 CRC64;
Query Match 54.8%; Score 40; DB 1; Length 304;
Best Local Similarity 50.0%; Pred. No. 79;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
QY 4 CRGDDSKVQE 13
DB 203 CRGGDARIHD 212
|||||:
RESULT 36
PNCB_ECO57
ID PNCB_ECO57 STANDARD; PRT; 399 AA.
AC Q8XDE8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Nicotinate phosphoribosyltransferase (EC 2.4.2.11) (NAPRTase).
GN Name=pncB; OrderedLocustNames=21279, EC61014;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,

RA Welch R.A., Blattner F.R.;
RT "genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7." ;
RL Nature 409:529-533 (2001).
PN [2]

```

1241
RP SEQUENCE FROM N.A.
RN STRAIN=O157:H7 / Sakai / RIMD 0509952 / EHEC;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayaehi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:1-12(2001).
CC -1- CATALYTIC ACTIVITY: Nicotinate D-ribonucleotide + diphosphate =
CC nicotinate + 5-phospho-alpha-D-ribose 1-diphosphate.
CC -1- PATHWAY: NAD biosynthesis; nicotinamide to NAMN; second step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the NAPRTase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).

```

[illegible]

Query Match 54.8%; Score 40; DB 1; Length 399;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 YRCRGDD 8
:|||||
Db 38 FRCRGDD 44

RESULT 37	
PNCB_ECOLI	
ID_PNCB_ECOLI	PRT; 399 AA.
STANDARD;	

AC	PI8133;
DT	01-NOV-1990 (Rel. 16, Created)
DT	01-MAY-1991 (Rel. 18, Last sequence update)
DT	25-JAN-2005 (Rel. 46, Last annotation update)
DE	Nicotinate phosphoribosyltransferase (EC 2.4.2.11) (NAPRTase) .
GN	NamespncB; OrderedLocusNames=b0931;
OS	Escherichia coli.
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC	Enterobacteriaceae; Escherichia.
OX	NCBI_TaxID=562;

RT and expression of the *pncB* gene encoding nicotinic acid phosphoribosyltransferase." ;
RL J. Biol. Chem. 265:17665-17672 (1990).
DN [2]

SEQUENCE FROM N.A.
STRAIN=K12 / MG1655;
MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
RA Blattner F., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of *Escherichia coli* K-12.";
RL science 277:1453-1474 (1997).

3]	SEQUENCE FROM N.A.
RP	STRAIN=K12;
RC	MEDLINE=970611202; PubMed=8905232;
RX	Oshima T., Aliba T., Baba T., Fujita K., Hayaashi K., Honjo A.,
RA	Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA	Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA	Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
RA	Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA	Yano M., Horiuchi T.;
RT	"A 718-kb DNA sequence of the <i>Escherichia coli</i> K-12 genome
RT	corresponding to the 12.7-28.0 min region on the linkage map.";
RL	DNA Res. 3:137-155(1996).
CC	-!- CATALYTIC ACTIVITY: Nicotinate D-ribonucleotide + diphosphate =
CC	nicotinate + 5-phospho-alpha-D-ribose 1-diphosphate.
CC	-!- PATHWAY: NAD biosynthesis; nicotinamide to NADN; second step.
CC	-!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC	-!- MISCELLANEOUS: Requires ATP as well as phosphoribosyl
CC	pyrophosphate for activity.
CC	-!- SIMILARITY: Belongs to the NAPRTase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC	EMBL; J05568; AAA24400.1; -
DR	EMBL; J05568; AAA24400.1; -
DR	EMBL; U00096; AAC74017.1; -
DR	EMBL; D90731; BAA35683.1; -
DR	EMBL; D90731; BAA35683.1; -
DR	PIR; JQ0756; JQ0756.

DR	EchoBASE; EB0735; -.
DR	Ecogene; EGI0742; pncB.
DR	HMAP; MF_00570; -; 1.
DR	InterPro; IPR007229; NAPTase.
DR	InterPro; IPR006406; Nic_Ppttrans.
DR	InterPro; IPR008967; P53_like_DNA_bnd.
DR	Pfam; PF04095; NAPTase; 1.
DR	PIRSF; PIRSF000484; Nicot_phos_ribo; 1.
DR	TIGRFAMs; TIGR01514; NAPTase; 1.
DR	Complete proteome; Direct protein sequencing; Glycosyltransferase;
KW	Pyridine nucleotide biosynthesis; Transferase.
KM	INIT MET 0 0
FT	SEQUENCE 399 AA; 45766 MW; 37F0C0D2C3BA5C45 CRC64;
SQ	

DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Nicotinate phosphoribosyltransferase (EC 2.4.2.11) (NAPRTase).
 GN Name=pncB; OrderedLocNames=STV1010, t1930;
 OS Salmonella typhi.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=601;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=CT18;
 RX MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
 RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Peltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
 RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
 RA Whitehead S., Barrrell B.G.;
 RA "Complete genome sequence of a multiple drug resistant Salmonella
 RT enterica serovar Typhi CT18";
 RL Nature 413:848-852(2001).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=Ty2 / ATCC 700931;
 RX MEDLINE=22531367; PubMed=12644504;
 RX DOI=10.1128/JB.185.7.2330-2337.2003;
 RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
 RA Burland V., Kodoyanni V., Schwartz D.C., Blattner F.R.;
 RA "Comparative genomics of Salmonella enterica serovar typhi strains Ty2
 RT and CT18";
 RL J. Bacteriol. 185:2330-2337(2003).
 CC -!- CATALYTIC ACTIVITY: Nicotinate D-ribonucleotide + diphosphate =
 CC nicotinate + 5-phospho-alpha-D-ribose 1-diphosphate.
 CC -!- PATHWAY: NAD biosynthesis; nicotinamide to NAMN; second step.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the NAPRTase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AL627268; CAD05404.1; -.
 DR EMBL; AE016840; AAO69546.1; -.
 DR HAMAP; MF_00570; -; 1.
 DR InterPro; IPR007229; NAPRTase.
 DR InterPro; IPR006406; Nic Prtrans.
 DR Pfam; PF04095; NAPRTase; 1.
 DR PIRSF; PIRSF000484; Nicot_phos_ribo; 1.
 DR TIGRFAMs; TIGR01514; NAPRTase; 1.
 DR TIGRFAMs; TIGR01514; Nicot_phos_ribo; 1.
 DR TIGRFAMs; TIGR01514; NAPRTase; 1.
 KW Complete proteome; Glycosyltransferase;
 KW Pyridine nucleotide biosynthesis; Transferase.
 FT INIT MET 0
 FT BY similarity.
 SQ SEQUENCE 399 AA; 45562 MW; ADF38AEAE4618214 CRC64;
 Query Match 54.8%; Score 40; DB 1; Length 399;
 Best Local Similarity 85.7%; Pred. No. 1.1e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 YRCRGDD 8
 Db :|||||
 38 FRCRGDD 44

RESULT 39
PNCB_SALTY

ID PNCB_SALTY STANDARD; PRT; 399 AA.
 AC F22253;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Nicotinate phosphoribosyltransferase (EC 2.4.2.11) (NAPRTase).
 GN Name=pncB; OrderedLocNames=STV1004;
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=602;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=91100340; PubMed=1987148;
 RA Vinitzky A., Teng H., Grubmeyer C.T.;
 RT "Cloning and nucleic acid sequence of the Salmonella typhimurium pncB
 RL gene and structure of nicotinate phosphoribosyltransferase.";
 RL J. Bacteriol. 173:536-540(1991).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RA "Complete genome sequence of Salmonella enterica serovar Typhimurium
 RT LT2";
 RL Nature 413:852-856(2001).
 CC -!- CATALYTIC ACTIVITY: Nicotinate D-ribonucleotide + diphosphate =
 CC nicotinate + 5-phospho-alpha-D-ribose 1-diphosphate.
 CC -!- PATHWAY: NAD biosynthesis; nicotinamide to NAMN; second step.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the NAPRTase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M55986; AAA27190.1; -.
 DR EMBL; AE008743; AAL19938.1; -.
 DR PIR; A39130; A39130.
 DR StyGene; SG10305; pncB.
 DR HAMAP; MF_00570; -; 1.
 DR InterPro; IPR007229; NAPRTase.
 DR InterPro; IPR006406; Nic Prtrans.
 DR InterPro; IPR008967; P53_like_DNA_bnd.
 DR Pfam; PF04095; NAPRTase; 1.
 DR PIRSF; PIRSF000484; Nicot_phos_ribo; 1.
 DR TIGRFAMs; TIGR01514; NAPRTase; 1.
 DR TIGRFAMs; TIGR01514; Nicot_phos_ribo; 1.
 KW Complete proteome; Glycosyltransferase;
 KW Pyridine nucleotide biosynthesis; Transferase.
 FT INIT MET 0
 FT BY similarity.
 SQ SEQUENCE 399 AA; 45530 MW; 90D9C3EDD8C092A6 CRC64;
 Query Match 54.8%; Score 40; DB 1; Length 399;
 Best Local Similarity 85.7%; Pred. No. 1.1e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 YRCRGDD 8
 Db :|||||
 38 FRCRGDD 44
 RESULT 40
 QYUD27
 ID QYUD27 PRELIMINARY; PRT; 400 AA.
 AC QYUD27;

DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nicotinate phosphoribosyltransferase.
GN Name=ncnB; OrderedLocusNames=S0992;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2457T;
RX MEDLINE=22590274; PubMed=12704152;
RX DOI=10.1128/IAI.71.5.2775-2786.2003;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Pournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RA "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T.";
RL Infect. Immun. 71:2775-2786(2003).
DR EMBL; AE016981; NAP16443.1; -
DR GO; GO:0004516; F:nicotinate phosphoribosyltransferase activity; IEA.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR GO; GO:0019357; P:nicotinate nucleotide biosynthesis; IEA.
DR GO; GO:0019363; P:pyridine nucleotide biosynthesis; IEA.
DR InterPro; IPR007229; NAPRTase.
DR InterPro; IPR006406; Nic_Prtase.
DR InterPro; IPR008967; P53_like_DNA_bnd.
DR Pfam; PF04095; NAPRTase; 1.
DR TIGRFAMs; TIGR01514; NAPRTase; 1.
DR TIGRFAMs; TIGR01514; Nicot_phos_ribo; 1.
DR Glycosyltransferase; Transferase.
SQ SEQUENCE 400 AA; 45911 MW; 53855BCB1A8703095 CRC64;
Query Match 54.8%; Score 40; DB 2; Length 400;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 YRCRGDD 8
Db :|||||
39 FRCRGDD 45
RESULT 41
Q83LN3 PRELIMINARY; PRT; 416 AA.
ID Q83LN3
AC Q83LN3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nicotinate phosphoribosyltransferase.
GN Name=ncnB; OrderedLocusNames=Sf0928;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).
DR EMBL; AE015122; AAN42557.1; -
DR GO; GO:0004516; F:nicotinate phosphoribosyltransferase activity; IEA.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR GO; GO:0019357; P:nicotinate nucleotide biosynthesis; IEA.

DR GO; GO:0019363; P:pyridine nucleotide biosynthesis; IEA.
DR InterPro; IPR007229; NAPRTase.
DR InterPro; IPR006406; Nic_Prtase.
DR InterPro; IPR008967; P53_like_DNA_bnd.
DR Pfam; PF04095; NAPRTase; 1.
DR PIRSF; PIRSF000484; Nicot_phos_ribo; 1.
DR TIGRFAMs; TIGR01514; NAPRTase; 1.
DR TIGRFAMs; TIGR01514; Nicot_phos_ribo; 1.
DR Glycosyltransferase; Transferase.
SQ SEQUENCE 416 AA; 47700 MW; E43D723542AF2DD5 CRC64;
Query Match 54.8%; Score 40; DB 2; Length 416;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 YRCRGDD 8
Db :|||||
55 FRCRGDD 61
RESULT 42
Q88J98 PRELIMINARY; PRT; 416 AA.
ID Q88J98
AC Q88J98;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nicotinate phosphoribosyltransferase (EC 2.4.2.11).
GN Name=ncnB; OrderedLocusNames=cl073;
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C6:HI / CFT073 / ATCC 700928 / UPEC;
RX MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252597999;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
DR EMBL; AE016758; AAN79541.1; -
DR GO; GO:0004516; F:nicotinate phosphoribosyltransferase activity; IEA.
DR GO; GO:0019357; P:nicotinate nucleotide biosynthesis; IEA.
DR GO; GO:0019363; P:pyridine nucleotide biosynthesis; IEA.
DR InterPro; IPR007229; NAPRTase.
DR InterPro; IPR006406; Nic_Prtase.
DR InterPro; IPR008967; P53_like_DNA_bnd.
DR Pfam; PF04095; NAPRTase; 1.
DR PIRSF; PIRSF000484; Nicot_phos_ribo; 1.
DR TIGRFAMs; TIGR01514; NAPRTase; 1.
DR TIGRFAMs; TIGR01514; Nicot_phos_ribo; 1.
DR Glycosyltransferase; Transferase.
SQ SEQUENCE 416 AA; 47722 MW; F00AFC07B04B1E23 CRC64;
Query Match 54.8%; Score 40; DB 2; Length 416;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 YRCRGDD 8
Db :|||||
55 FRCRGDD 61
RESULT 43
Q04B CAEL STANDARD; PRT; 577 AA.
ID Q04B CAEL
AC Q17426;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Hypothetical pseudouridine synthase in chromosome V (EC 4.2.1.70).

GN ORPNames=B0024.11;
 OC Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RG The C. elegans sequencing consortium;
 RT "Genome sequence of the nematode C. elegans: a platform for
 investigating biology."
 RL Science 282:2012-2018(1998).
 CC -1- CATALYTIC ACTIVITY: Uracil + D-ribose 5-phosphate = pseudouridine
 CC 5'-phosphate + H(2)O.
 CC -1- SIMILARITY: Belongs to the pseudouridine synthase truD family.
 CC -1- SIMILARITY: Contains 1 TRUD domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z71178; CAA94883.1; -.
 DR PIR; T18646; T18646.
 DR WormBase; WBGenome0007101; B0024.11.
 DR WormPep; B0024.11; CE05155.
 DR InterPro; IPR001656; PaU_synth_TRUD.
 DR Pfam; PF01142; Trud; 1.
 DR PROSITE; PS50984; TRUD; 1.
 DR PROSITE; PS01268; UPF0024; 1.
 DR DOMAIN 265 472
 KW Hypothetical protein; Lyase; TRUD.
 FT DOMAIN 265 472
 FT ACT_SITE 188 188 By similarity.
 SQ SEQUENCE 577 AA; 64631 MW; A1F3BC06481CD281 CRC64;

Query Match 54.8%; Score 40; DB 1; Length 577;
 Best Local Similarity 46.2%; Pred. No. 1.6e+02;
 Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 NYRCRGDDSKVQ 13
 :: ||||| ::
 Db 544 SFKTRGDDKTEE 556

RESULT 44
 Q62YB1 PRELIMINARY; PRT; 586 AA.
 ID Q62YB1
 AC Q62YB1; 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Alpha amylase, catalytic subdomain.
 GN Names=yvdf; ORFNames=BL00497;
 OS Bacillus licheniformis DSM 13.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=279010;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 14590;
 RX MEDLINE=20365717; PubMed=10910347; DOI=10.1038/35018003;
 RA Simpson A.J.G., Reinach P.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvares R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
 RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hohnsels J.D., Junqueira M.L., Kemper E.L., Leitao J.P.,
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,

Query Match 54.8%; Score 40; DB 2; Length 586;
 Best Local Similarity 70.0%; Pred. No. 1.6e+02;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 RCRGDDSKVQ 12
 ||||| ::
 Db 431 RCGDDKKVR 440

RESULT 45
 Q65MW8 PRELIMINARY; PRT; 587 AA.
 ID Q65MW8
 AC Q65MW8; 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Maltogenic alpha-amylase (EC 3.2.1.133).
 GN ORFNames=BLI00658;
 OS Bacillus licheniformis DSM 13.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=279010;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 13;
 RX PubMed=15383718;
 RA Veith B., Herzberg C., Steckel S., Feesche J., Maurer K.H.,
 RA Ehrenreich P., Baumer S., Henne A., Liesegang H., Merkl R.,
 RA Ehrenreich A., Gottschalk G.;
 RT "The Complete Genome Sequence of Bacillus licheniformis DSM13, an
 RT Organism with Great Industrial Potential.";
 RL J. Mol. Microbiol. Biotechnol. 7:204-211(2004).
 DR EMBL; AE017333; AAU39596.1; -.
 KW Glycosidase; Hydrolase.
 SQ SEQUENCE 587 AA; 67461 MW; 7F3C26DA8C994309 CRC64;

Query Match 54.8%; Score 40; DB 2; Length 587;
 Best Local Similarity 70.0%; Pred. No. 1.6e+02;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 RCRGDDSKVQ 12
 ||||| ::
 Db 432 RCGDDKKVR 441

RESULT 46
 Q9PA43 PRELIMINARY; PRT; 721 AA.
 ID Q9PA43
 AC Q9PA43; 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE L-ascorbate oxidase.
 GN OrderedLocusNames=Xf2677;
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xylella.
 OX NCBI_TaxID=2371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9a5C;
 RX MEDLINE=20365717; PubMed=10910347; DOI=10.1038/35018003;
 RA Simpson A.J.G., Reinach P.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvares R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
 RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hohnsels J.D., Junqueira M.L., Kemper E.L., Leitao J.P.,
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,

RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuchioka M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Weidman J., Setubal J.C.,
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
DR EMBL; AB004073; AAF85474.1; -.
DR FIR; H82528; H82528.
DR HSSP; Q51883; 1BAW.
DR GO; GO:0005507; F:copper ion binding; IEA.
DR InterPro; IPR001117; Cu-oxidase.
DR InterPro; IPR008972; Cupredoxin.
DR InterPro; IPR002355; Cu,ox.copper_BS.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 1.
DR PROSITE; PS00080; MULTICOPPER_OXIDASE2; 1.
KW Complete proteome.
SQ SEQUENCE 721 AA; 80882 MW; 847E513614FB1955 CRC64;

Query Match 54.8%; Score 40; DB 2; Length 721;
Best Local Similarity 54.5%; Pred. No. 2e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 YRCRGDSKVQ 12
DB 282 YACRGNDGKLG 292

RESULT 47
Q7UKD4 PRELIMINARY; PRT; 1129 AA.
ID Q7UKD4
AC Q7UKD4;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to serine/threonine protein kinase related protein-putative
DE FQO-dependent oxidoreductase.
GN OrderedLocusNames=RB10706;
OS Rhodospirillum rubrum
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RP [1]
RC SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR EMBL; BX294152; CAD76947.1; -.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR002372; PQO repeat.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR SMART; SM00564; PQO; 3.
KW Complete proteome; Kinase; Serine/threonine-protein kinase.
SQ SEQUENCE 1129 AA; 122926 MW; A5AC5DF164BE06F1 CRC64;

Query Match 54.8%; Score 40; DB 2; Length 1129;

Best Local Similarity 60.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 RCRGDSKVQ 12
DB 500 RCRGEDGEVE 509

RESULT 48
Q7RQM4 PRELIMINARY; PRT; 1615 AA.
ID Q7RQM4;
AC Q7RQM4;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Multidomain scavenger receptor protein PDSR precursor.
GN Name=PY01071;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=73239;
RV [1]
RP SEQUENCE FROM N.A.
RC STRAIN=17XNL;
RX PubMed=12368865; DOI=10.1038/nature01099;
RA Carlton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Perteau M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA Shalom S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoab A., Cummings L.M.,
RA Florens L., Yates F. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria
RT parasite Plasmodium yoelii yoelii.";
RL Nature 419:512-519(2002).
CC -!- SIMILARITY: Contains 1 PLAT domain.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AABL01000281; EAA20072.1; -.
DR HSSP; O57811; 1G8A.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0005044; F:scavenger receptor activity; IEA.
DR GO; GO:0006364; P:rRNA processing; IEA.
DR InterPro; IPR008985; ConA like lec_gl.
DR InterPro; IPR000692; Fibrillarin.
DR InterPro; IPR004043; LCCL.
DR InterPro; IPR001024; Lipoxigenase_LH2.
DR InterPro; IPR008976; PLAT_LH2.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF01269; Fibrillarin; 1.
DR Pfam; PF03815; LCCL; 4.
DR Pfam; PF01477; PLAT; 1.
DR Pfam; PF00530; SRCR; 2.
DR PRINTS; PD00052; FIBRILLARIN.
DR ProDom; PD004637; Fibrillarin; 1.
DR PROSITE; PS00566; FIBRILLARIN; 1.
DR PROSITE; PS00820; LCCL; 4.
DR PROSITE; PS00095; PLAT; 1.
DR PROSITE; PS0287; SRCR_2; 2.
KW Receptor.
SQ SEQUENCE 1615 AA; 180943 MW; 94053C4C9AC93505 CRC64;

Query Match 54.8%; Score 40; DB 2; Length 1615;
Best Local Similarity 41.7%; Pred. No. 4.8e+02;
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYRCRGDSKVQ 12

```
Db 1:||||: 614 NFRCKGDEPNLK 625

RESULT 49
ID Q848W4 PRELIMINARY; PRT; 46 AA.
AC Q848W4
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Bacillus megaterium.
OC Plasmid pBM400.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1404;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OM B1551;
RX MEDLINE=22964369; PubMed=14602653;
DOI=10.1128/AEM.69.11.6888-6898.2003;
RA Scholle M.D., White C.A., Kunitalalyaan M., Vary P.S.;
RT "Sequencing and characterization of pBM400 from Bacillus megaterium QM B1551."
RL Appl. Environ. Microbiol. 69:6888-6898(2003).
DR EMBL; AF142677; AA052802.1; -.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 46 AA; 5449 MW; F4B4C4A3A050B138 CRC64;

Query Match 53.4%; Score 39; DB 2; Length 46;
Best Local Similarity 60.0%; Pred. No. 16;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYRCRGDSDK 10
Db |||||:
35 NYRCKGVDK 44

RESULT 50
Q8BUS2 PRELIMINARY; PRT; 154 AA.
AC Q8BUS2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Mus musculus ES cells cDNA, RIKEN full-length enriched library,
DE clone:C330016E03 product:SRCRB-S4D PROTEIN homolog (Fragment).
GN Name=Srcrb4d;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning."
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
```

```
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes."
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsumai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Hazada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer."
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Aono H., Arai A.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hirooka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yamanaka I., Yasunishi A.,
RA Yoshida K., Yoshino M., Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK082786; BAC38619.1; -.
DR HSSP; Q08380; 1BY2.
DR MGD; MGI:1924709; Scrb4d.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005044; F:scavenger receptor activity; IEA.
DR InterPro; IPR001190; Scrb_receptor.
DR Pfam; PF00530; SRCR; 1.
DR PRINTS; PR00258; SPERACTRCPTR.
DR SMART; SM00202; SR; 1.
DR PROSITE; PS00420; SRCR_1; 1.
DR PROSITE; PS0287; SRCR_2; 1.
DR NON_TER 1
FT SEQUENCE 154 AA; 17028 MW; 8932AA223E2F14F5 CRC64;

Query Match 53.4%; Score 39; DB 2; Length 154;
Best Local Similarity 66.7%; Pred. No. 58;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDSDS 9
Db |||||:
118 NYRCRGDES 126

Search completed: September 7, 2005, 19:57:33
Job time : 51.6512 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2005, 19:49:10 ; Search time 7.67442 Seconds
(without alignments)
125.373 Million cell updates/sec

Title: US-10-812-238B-20

Perfect score: 59

Sequence: 1 NYRCRGDDSK 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database :

1: PIR 79:*

2: PIR1:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	67.8	150	2 G83840	ferric uptake regu
2	40	67.8	400	1 A39130	nicotinate phospho
3	40	67.8	400	1 JQ0756	nicotinate phospho
4	40	67.8	400	2 AB0617	nicotinate phospho
5	40	67.8	400	2 F90755	nicotinate phospho
6	40	67.8	400	2 D85619	nicotinate phospho
7	40	67.8	705	2 A70669	probable acid-CoA
8	39	66.1	179	2 T22521	hypothetical prote
9	38	64.4	389	2 T27085	hypothetical prote
10	38	64.4	721	2 H82528	L-ascorbate oxidas
11	38	64.4	1267	1 MWXR31	lambda 3 protein -
12	38	64.4	1267	1 MWXR33	lambda 3 protein -
13	38	62.7	430	2 S75313	serine-tRNA ligase
14	37	62.7	504	2 C64398	hypothetical prote
15	37	62.7	1268	2 T31420	C-terminal domain-
16	37	62.7	1268	2 T31420	probable membrane
17	36	61.0	347	2 T35013	4-hydroxyphenylace
18	36	61.0	461	1 E69360	hypothetical prote
19	36	61.0	470	2 E72376	ABC transporter-as
20	36	61.0	507	2 AD3382	breast/ovarian can
21	36	61.0	1812	2 I49350	hypothetical prote
22	35	59.3	138	2 C90460	leukocyte activati
23	35	59.3	187	2 A10520	Amp protein (anhy
24	35	59.3	269	2 A46506	hypothetical prote
25	35	59.3	274	2 T32736	hypothetical prote
26	35	59.3	329	2 T20546	probable chitinase
27	35	59.3	453	2 T24127	anthranilate synth
28	35	59.3	513	2 B87484	serine/threonine-s
29	35	59.3	656	2 T10664	

30	35	59.3	662	2 I37892	IL12 receptor comp
31	35	59.3	1323	2 T37533	coronin-like prote
32	34.5	58.5	325	2 B81364	probable periplasm
33	34	57.6	112	2 H70232	hypothetical prote
34	34	57.6	122	2 T04118	mitochondrial proc
35	34	57.6	172	2 F81439	translation initia
36	34	57.6	180	2 AC3270	hypothetical expor
37	34	57.6	255	2 G84320	hypothetical prote
38	34	57.6	259	2 E84427	hypothetical prote
39	34	57.6	270	2 C96731	unknown protein F5
40	34	57.6	273	2 T10610	hypothetical prote
41	34	57.6	296	2 A95408	hypothetical prote
42	34	57.6	344	2 T05437	hypothetical prote
43	34	57.6	348	2 H85256	hypothetical prote
44	34	57.6	375	2 B96551	hypothetical prote
45	34	57.6	393	2 B86189	protein T25N20.9 [
46	34	57.6	407	2 C36786	hypothetical prote
47	34	57.6	413	2 F96743	probable C2H2-type
48	34	57.6	482	2 H71400	hypothetical prote
49	34	57.6	596	2 B41627	furin (EC 3.4.21.7
50	34	57.6	783	2 A41627	furin (EC 3.4.21.7
51	34	57.6	801	2 A89862	Na+/H+ antiporter
52	34	57.6	850	2 S54553	SIN3 protein-bindi
53	34	57.6	988	2 F86315	protein T10022.13
54	34	57.6	1030	2 F96763	hypothetical prote
55	34	57.6	1036	2 T18530	phospholipase D (E
56	33	55.9	152	2 A10448	probable ribonucle
57	33	55.9	167	2 B84536	hypothetical prote
58	33	55.9	185	2 H87107	ribosome recycling
59	33	55.9	200	2 T02826	ribosomal protein
60	33	55.9	214	2 S46476	cysteine proteinas
61	33	55.9	226	2 AD2198	hypothetical prote
62	33	55.9	238	2 T11718	hypothetical prote
63	33	55.9	248	2 F64908	probable dehydroge
64	33	55.9	248	2 D90897	probable oxidoredu
65	33	55.9	248	2 C85720	probable oxidoredu
66	33	55.9	273	2 D81878	probable oxidoredu
67	33	55.9	336	2 G82228	formiminoglutamase
68	33	55.9	352	2 C96643	hypothetical prote
69	33	55.9	370	2 F84550	hypothetical prote
70	33	55.9	389	1 A48329	histidinol-phospha
71	33	55.9	421	2 A83467	glutamate dehydrog
72	33	55.9	489	2 C87480	conserved hypothet
73	33	55.9	521	2 T18896	glypican 1 precurs
74	33	55.9	569	2 D85354	hypothetical prote
75	33	55.9	577	2 T18646	hypothetical prote
76	33	55.9	600	2 E87665	arginyl-tRNA synth
77	33	55.9	675	2 S74399	sensory transducti
78	33	55.9	680	2 G82526	ATP-dependent DNA
79	33	55.9	707	2 F86925	probable acyl-CoA
80	33	55.9	725	2 JE0100	neural cell adhesi
81	33	55.9	730	2 JH0798	fasciclin IV precu
82	33	55.9	842	2 C83177	probable phosphotr
83	33	55.9	955	2 T18435	hypothetical prote
84	33	55.9	1035	2 T42093	phospholipase D (E
85	33	55.9	1036	2 T13732	phospholipase D (E
86	33	55.9	1037	2 T13943	phospholipase D (E
87	33	55.9	1041	2 PQ0442	polyprotein - barl
88	33	55.9	1074	2 T17203	phospholipase (EC
89	33	55.9	1074	2 T13725	phospholipase D (E
90	33	55.9	1075	2 T46635	phospholipase D (E
91	33	55.9	1092	1 UN0635	neural cell adhesi
92	33	55.9	1202	2 PQ0440	polyprotein - barl
93	33	55.9	1231	1 NBH0H	complement factor
94	33	55.9	1470	2 S45323	genome polyprotein
95	33	55.9	2212	2 T28157	erythrocyte membra
96	33	55.9	2410	1 JQ1948	genome polyprotein
97	33	55.9	2412	1 JQ1537	hypothetical prote
98	33	55.9	2647	2 T28161	LDL-receptor-relat
99	33	55.9	4753	1 A47437	syngingomycin synth
100	33	55.9	9376	2 T14593	probable sulfatase
101	32.5	55.1	571	2 A10506	DNA-binding protei
102	32	54.2	66	1 DNVFB	

103	32	54.2	74	1	DNPVAS	DNA-binding protei	176	31	52.5	129	2	S48902	hypothetical prote
104	32	54.2	97	2	A82158	hypothetical prote	177	31	52.5	145	2	F82133	hypothetical prote
105	32	54.2	111	2	C71401	hypothetical prote	178	31	52.5	147	2	T18586	hypothetical prote
106	32	54.2	129	2	AF0527	conserved hypothet	179	31	52.5	161	2	AC0883	conserved hypothet
107	32	54.2	145	2	A83010	conserved hypothet	180	31	52.5	170	2	AG2479	aminoglycoside N6'
108	32	54.2	154	2	A60998	replication protei	181	31	52.5	182	2	A25468	T-cell surface gly
109	32	54.2	158	2	AB0426	transcription elon	182	31	52.5	222	2	S26994	cellulose 1,4-beta
110	32	54.2	161	1	TPHUCC	troponin C, cardia	183	31	52.5	228	2	S26995	cellulose 1,4-beta
111	32	54.2	161	1	TPBOCC	troponin C, cardia	184	31	52.5	228	2	T47425	NAC domain-like pr
112	32	54.2	161	1	TPBRCW	troponin C, cardia	185	31	52.5	236	2	S26993	cellulose 1,4-beta
113	32	54.2	161	2	S07450	troponin C - quail	186	31	52.5	240	2	E86418	probable eukaryoti
114	32	54.2	161	2	A27204	troponin C, cardia	187	31	52.5	245	2	G82175	hypothetical prote
115	32	54.2	161	2	JU0035	troponin C, cardia	188	31	52.5	268	2	T31699	hypothetical prote
116	32	54.2	161	2	JW0064	slow cardiac tropo	189	31	52.5	274	2	T10270	protein kinase (EC
117	32	54.2	161	2	A32620	troponin C, cardia	190	31	52.5	280	2	T43011	suppressor protein
118	32	54.2	171	1	RWHD1	T-cell surface gly	191	31	52.5	284	2	AG3231	conjugal transfer
119	32	54.2	183	2	AF0904	probable exported	192	31	52.5	285	2	G85016	probable myb-relat
120	32	54.2	190	2	T46053	ADP-ribosylation f	193	31	52.5	286	2	S50855	neurotrophin-6 - s
121	32	54.2	204	2	E84173	hypothetical prote	194	31	52.5	292	2	T10106	chitinase (EC 3.2.
122	32	54.2	221	2	T26921	hypothetical prote	195	31	52.5	314	2	C90256	conserved hypothet
123	32	54.2	229	2	I46614	nerve growth facto	196	31	52.5	317	2	T18695	hypothetical prote
124	32	54.2	241	2	JL0097	nerve growth facto	197	31	52.5	323	2	T27640	hypothetical prote
125	32	54.2	245	2	I56570	beta-nerve growth	198	31	52.5	338	2	H83228	hypothetical prote
126	32	54.2	289	1	NGHUBM	nerve growth facto	199	31	52.5	344	2	T27664	hypothetical prote
127	32	54.2	289	2	T44599	oligopeptide trans	200	31	52.5	346	2	T49085	hypothetical prote
128	32	54.2	303	1	NGRTBA	nerve growth facto	201	31	52.5	347	2	JC7828	glucuronyltransfer
129	32	54.2	305	2	E84261	glycosyl transfera	202	31	52.5	359	2	AF3184	beta-lactamase (im
130	32	54.2	307	1	NGMSMG	nerve growth facto	203	31	52.5	370	2	T22082	hypothetical prote
131	32	54.2	324	2	H86244	lysophospholipase	204	31	52.5	422	2	S34199	polygalacturonase
132	32	54.2	350	2	T25451	transforming growt	205	31	52.5	435	2	T32195	hypothetical prote
133	32	54.2	357	2	AG1072	conserved hypothet	206	31	52.5	460	2	JC4214	translation elonga
134	32	54.2	363	2	G95937	probable mureinpep	207	31	52.5	461	2	T33747	hypothetical prote
135	32	54.2	383	2	A96704	hypothetical prote	208	31	52.5	497	2	B97555	hypothetical prote
136	32	54.2	387	2	JL0101	methane monooxygen	209	31	52.5	497	2	AD2775	fumarate hydratase
137	32	54.2	427	2	G02034	Killer cell inhibi	210	31	52.5	501	2	T37747	hypothetical prote
138	32	54.2	435	2	G69194	L-asparaginase I -	211	31	52.5	510	2	S41943	cellulose 1,4-beta
139	32	54.2	444	2	G01924	KIR (cl-2) NK rece	212	31	52.5	511	2	S44716	cellulose 1,4-beta
140	32	54.2	444	2	G01925	KIR (cl-11) NK rec	213	31	52.5	530	2	S65773	glutamyl-tRNA redu
141	32	54.2	455	2	G01923	KIR (cl-5) NK rece	214	31	52.5	530	2	G86233	hypothetical prote
142	32	54.2	457	2	T48036	hypothetical prote	215	31	52.5	540	2	S41942	cellulose 1,4-beta
143	32	54.2	462	2	H95914	hypothetical membr	216	31	52.5	542	2	T10245	glutamyl-tRNA redu
144	32	54.2	474	1	OMHU1B	alpha-1-B-glycopro	217	31	52.5	548	2	T22137	hypothetical prote
145	32	54.2	486	2	A83284	probable oxidoredu	218	31	52.5	551	2	S58238	HR38 protein - fr
146	32	54.2	502	2	S38139	LAS1 protein - yea	219	31	52.5	557	2	T16696	hypothetical prote
147	32	54.2	505	2	T19106	probable serine ca	220	31	52.5	567	2	E95901	conserved hypothet
148	32	54.2	509	2	T02864	probable Zn finger	221	31	52.5	576	2	T41587	probable carbon ca
149	32	54.2	516	2	D96682	protein F1822.18 [222	31	52.5	577	2	T02401	probable beta-gluc
150	32	54.2	536	2	T20736	hypothetical prote	223	31	52.5	587	2	F84475	hypothetical prote
151	32	54.2	539	2	A54294	cell division cont	224	31	52.5	594	2	I58386	receptor tyrosine
152	32	54.2	562	2	AF0852	secretory protein i	225	31	52.5	597	2	E85090	probable transposo
153	32	54.2	563	2	S54420	invasion protein i	226	31	52.5	612	2	T02414	probable protein k
154	32	54.2	589	2	F70033	glucan 1,4-alpha-m	227	31	52.5	615	2	S54468	AIPI protein - yea
155	32	54.2	607	2	S77092	hypothetical prote	228	31	52.5	629	2	H86383	probable wall-asso
156	32	54.2	664	2	E70770	probable atp-depen	229	31	52.5	651	2	F90111	DNA primase (impor
157	32	54.2	729	2	T45951	hypothetical prote	230	31	52.5	662	2	T41442	omnipotent nonsens
158	32	54.2	731	2	T29129	probable ATP/GTP b	231	31	52.5	662	2	T51948	omnipotent nonsens
159	32	54.2	864	2	T30441	probable capsid-as	232	31	52.5	710	1	I51283	hepatocyte growth
160	32	54.2	1095	2	T00329	hypothetical prote	233	31	52.5	728	1	S59553	arginine decarboxy
161	32	54.2	1116	2	T38073	serine/threonine-p	234	31	52.5	729	2	T23972	hypothetical prote
162	32	54.2	1283	2	T13799	neurexin iv - frui	235	31	52.5	778	2	B70667	hypothetical prote
163	32	54.2	1331	2	S05011	calcium channel al	236	31	52.5	838	2	T20125	hypothetical prote
164	32	54.2	1429	2	JU0609	nitric-oxide synth	237	31	52.5	878	2	H83184	phosphoenolpyruvat
165	32	54.2	1429	2	S16233	nitric-oxide synth	238	31	52.5	888	2	H88085	protein TiiF1.8 li
166	32	54.2	1433	2	G01946	nitric-oxide synth	239	31	52.5	1150	2	T18303	DNA binding regula
167	32	54.2	2042	2	T18399	variant-specific s	240	31	52.5	1162	2	T18199	reverse transcript
168	32	54.2	2166	2	S11339	calcium channel pr	241	31	52.5	1330	2	S49010	embryonic receptor
169	32	54.2	2171	2	S05054	calcium channel al	242	31	52.5	1333	2	I78875	receptor tyrosine
170	32	54.2	2204	1	RRNZNV	genome polyprotein	243	31	52.5	1336	2	I60598	Fit-1 tyrosine kin
171	32	54.2	2657	2	T18497	hypothetical prote	244	31	52.5	1364	2	T00250	MEGF2 protein - hu
172	32	54.2	2911	2	T20566	hypothetical prote	245	31	52.5	1390	2	S51364	sperm tail-specifi
173	32	54.2	3228	2	T21381	hypothetical prote	246	31	52.5	1475	2	T29809	hypothetical prote
174	31.5	53.4	189	2	G96514	hypothetical prote	247	31	52.5	1489	2	T38842	probable RAS GTPas
175	31	52.5	125	2	T32471	hypothetical prote	248	31	52.5	1580	2	T26204	hypothetical prote

249	31	52.5	1750	2	H64403	ribonucleoside-tri	322	30	50.8	376	2	T32837	hypothetical prote
250	31	52.5	1763	2	T19183	hypothetical prote	323	30	50.8	377	2	B72275	probable aspartate
251	31	52.5	1769	2	T19184	hypothetical prote	324	30	50.8	382	2	G85234	hypothetical prote
252	31	52.5	2051	2	T30938	receptor tyrosine	325	30	50.8	384	1	S41584	histidinol-phospha
253	31	52.5	3328	2	T30835	breast cancer tumo	326	30	50.8	385	1	S48456	hypothetical prote
254	31	52.5	3329	2	T32905	breast cancer susc	327	30	50.8	392	2	A86252	probable nicotinat
255	31	52.5	3329	2	T30904	breast cancer tumo	328	30	50.8	401	2	AG0172	hypothetical prote
256	31	52.5	3432	1	GNWVJS	genome polyprotein	329	30	50.8	408	2	D84635	hypothetical prote
257	31	52.5	3432	1	GNWVJE	genome polyprotein	330	30	50.8	414	2	T24563	hypothetical prote
258	31	52.5	4549	2	T20771	hypothetical prote	331	30	50.8	416	2	B55649	TNFR-associated pr
259	31	52.5	4667	2	T20774	hypothetical prote	332	30	50.8	423	2	B84751	hypothetical prote
260	30.5	51.7	732	2	T16422	hypothetical prote	333	30	50.8	428	2	S10013	alpha-amylase (EC
261	30.5	51.7	1140	2	B97820	hypothetical prote	334	30	50.8	434	1	DEEBHT	histidinol dehydro
262	30	50.8	36	2	A06002	annexin II - golde	335	30	50.8	434	2	AD0764	histidinol dehydro
263	30	50.8	39	2	A36453	decorsin - leech (336	30	50.8	434	2	S12775	alpha-amylase (EC
264	30	50.8	57	1	T1FHPB	proteinase inhibit	337	30	50.8	462	2	AH1130	glutamyl-decarbox
265	30	50.8	85	2	T10571	hypothetical prote	338	30	50.8	465	2	T05734	glutamyl-tRNA redu
266	30	50.8	89	2	B91173	phosphotransferase	339	30	50.8	465	2	T00090	glycosyltransferas
267	30	50.8	89	2	B86019	probable phosphoca	340	30	50.8	467	2	A47388	serine/threonine p
268	30	50.8	112	2	B86263	F13K23.19 protein	341	30	50.8	468	2	T10595	hypothetical prote
269	30	50.8	116	2	H75591	hypothetical prote	342	30	50.8	470	2	F85164	hypothetical prote
270	30	50.8	133	2	I45927	membrane-bound inm	343	30	50.8	470	2	S54089	hypothetical prote
271	30	50.8	135	2	I49275	protein kinase Sry	344	30	50.8	483	2	A39676	protein kinase Sry
272	30	50.8	141	2	A40463	integrin alpha-6 c	345	30	50.8	484	2	E64432	spore coat polysac
273	30	50.8	147	1	LZDK	lysozyme (EC 3.2.1	346	30	50.8	485	2	T32144	hypothetical prote
274	30	50.8	149	2	B40463	integrin alpha-6 c	347	30	50.8	491	2	E87452	ankyrin-related pr
275	30	50.8	163	2	D83883	hypothetical prote	348	30	50.8	499	2	H83148	AMP nucleosidase p
276	30	50.8	183	2	S05569	signaling protein	349	30	50.8	503	2	T46852	anthranilate synth
277	30	50.8	183	2	B90643	Ampb protein [impo	350	30	50.8	514	2	F64899	nitrate reductase
278	30	50.8	183	2	B85494	regulates ampC [im	351	30	50.8	514	2	C85730	cryptic nitrate re
279	30	50.8	183	2	C65110	hypothetical prote	352	30	50.8	514	2	F90887	cryptic nitrate re
280	30	50.8	183	2	C85983	hypothetical prote	353	30	50.8	528	2	T05732	probable glutamyl-
281	30	50.8	183	2	H91137	hypothetical prote	354	30	50.8	529	2	S18453	variant surface gl
282	30	50.8	188	2	H90255	adenylate kinase r	355	30	50.8	543	2	E96616	hypothetical prote
283	30	50.8	195	2	B85364	hypothetical prote	356	30	50.8	546	1	S13757	RNA helicase DBP2
284	30	50.8	211	2	S55129	transcription fact	357	30	50.8	548	2	T28910	hypothetical prote
285	30	50.8	219	2	AH2353	conserved hypothet	358	30	50.8	549	2	H96944	arginine degradati
286	30	50.8	226	2	H84237	hypothetical prote	359	30	50.8	552	2	T10186	glutamyl-tRNA redu
287	30	50.8	231	2	T32146	hypothetical prote	360	30	50.8	566	2	F84903	probable ubiquitin
288	30	50.8	254	2	E98329	hypothetical prote	361	30	50.8	573	2	S00661	hypothetical prote
289	30	50.8	255	2	A84723	hypothetical prote	362	30	50.8	573	2	JC4335	anti-mullerian hor
290	30	50.8	264	2	T47183	hypothetical prote	363	30	50.8	586	2	T20353	hypothetical prote
291	30	50.8	266	2	T10609	hypothetical prote	364	30	50.8	602	1	S46676	probable phosphoe
292	30	50.8	270	1	ICMS1	interleukin-1 alph	365	30	50.8	633	2	T04835	probable serine/th
293	30	50.8	270	1	JX0064	interleukin-1 alph	366	30	50.8	634	2	S77096	icfG protein sir18
294	30	50.8	270	1	F86177	protein F19p19.1 [367	30	50.8	665	2	E86775	ribonuclease [impo
295	30	50.8	273	2	H81140	oxidoreductase, sh	368	30	50.8	673	2	S36336	probable transcrip
296	30	50.8	285	2	A86419	probable eukaryoti	369	30	50.8	738	1	TFHUM	melanotransferrin
297	30	50.8	288	2	C83518	hypothetical prote	370	30	50.8	738	2	T27289	hypothetical prote
298	30	50.8	297	2	H44020	hypothetical prote	371	30	50.8	752	1	PIIVBC	RNA-directed RNA p
299	30	50.8	303	2	B84744	hypothetical prote	372	30	50.8	752	1	PIIVBL	RNA-directed RNA p
300	30	50.8	306	2	I49068	protein kinase Sry	373	30	50.8	752	1	PIIVBW	RNA-directed RNA p
301	30	50.8	309	1	KIECRB	ribokinase (EC 2.7	374	30	50.8	830	2	T04848	protein kinase hom
302	30	50.8	309	2	F91215	ribokinase [import	375	30	50.8	862	2	I49583	differentiation an
303	30	50.8	309	2	G86061	ribokinase [import	376	30	50.8	868	2	A46512	CD22 homolog/B lym
304	30	50.8	312	2	T35940	probable transport	377	30	50.8	931	2	AF3276	ATP-dependent clp
305	30	50.8	315	2	T10613	hypothetical prote	378	30	50.8	932	2	F84465	hypothetical prote
306	30	50.8	321	2	S13550	anther-specific pr	379	30	50.8	964	2	E71460	probable outer mem
307	30	50.8	326	2	AE2466	acetyl-CoA carboxy	380	30	50.8	976	2	H81722	polymorphic membra
308	30	50.8	336	2	E69845	thiamin biosynthes	381	30	50.8	1030	2	C86189	unknown protein. 2
309	30	50.8	339	1	LUBO36	annexin II - bovin	382	30	50.8	1048	2	C86189	protein T55N20.11
310	30	50.8	339	1	LUBO36	annexin II - human	383	30	50.8	1073	2	B36429	integrin alpha-6 c
311	30	50.8	339	1	LUMS36	annexin II - mouse	384	30	50.8	1091	2	A41543	integrin alpha-6 c
312	30	50.8	339	2	S33700	annexin II - rat	385	30	50.8	1132	2	A35089	myosin-binding pro
313	30	50.8	339	2	B34895	transcription fact	386	30	50.8	1205	2	T18517	procollagen N-endo
314	30	50.8	341	2	S55277	annexin II - rat	387	30	50.8	1207	2	T23754	hypothetical prote
315	30	50.8	342	2	T45577	protein kinase-lik	388	30	50.8	1338	2	A49634	aldehyde oxidase (
316	30	50.8	344	1	TXWL3	transcription fact	389	30	50.8	1372	2	T25933	hypothetical prote
317	30	50.8	352	2	C98170	hypothetical prote	390	30	50.8	1375	2	T13822	frazzled gene prot
318	30	50.8	373	2	T47115	probable 4-carboxy	391	30	50.8	1385	2	S34230	156K protein - Pla
319	30	50.8	375	2	I60125	PDGF receptor beta	392	30	50.8	1419	2	A41182	collagen alpha 1(I
320	30	50.8	375	2	AC1350	N-acetylmuramoyl-L	393	30	50.8	1433	1	GNVUBW	M polyprotein prec
321	30	50.8	375	2	AF1720	N-acetylmuramoyl-L	394	30	50.8	1526	2	T13823	frazzled gene prot

395	30	50.8	1535	2	T49042	hypothetical prote	468	29	49.2	243	2	A26311	nerve growth facto
396	30	50.8	1806	2	T23298	hypothetical prote	469	29	49.2	243	2	I51193	nerve growth facto
397	30	50.8	1847	2	T13308	probable vitellog	470	29	49.2	244	1	VCCVSV	coat protein - mal
398	30	50.8	2515	2	A41519	posterior-group pr	471	29	49.2	244	2	T10112	capsid protein - m
399	30	50.8	3005	2	S33642	homeotic protein z	472	29	49.2	247	2	T02253	chlorophyll a/b-bi
400	30	50.8	3433	1	GNWVKV	genome polyprotein	473	29	49.2	249	2	A13401	arginyltransferase
401	30	50.8	3434	1	GNWVMV	genome polyprotein	474	29	49.2	252	2	C83837	hypothetical prote
402	30	50.8	3924	2	S37431	ankyrin 2, neurona	475	29	49.2	252	2	T46247	hypothetical prote
403	30	50.8	4543	1	A53102	alpha-2-macroglobu	476	29	49.2	264	2	F91017	hypothetical prote
404	30	50.8	4544	1	S02392	alpha-2-macroglobu	477	29	49.2	264	2	H85861	hypothetical prote
405	30	50.8	4545	1	S25111	alpha-2-macroglobu	478	29	49.2	265	2	D83577	conserved hypotet
406	30	50.8	4660	2	T42737	gp330 protein prec	479	29	49.2	270	2	G96686	hypothetical prote
407	30	50.8	4687	1	A39638	plectin - rat	480	29	49.2	271	2	H97230	alpha/beta superfa
408	29.5	50.0	83	2	T23366	hypothetical prote	481	29	49.2	273	2	T39405	hypothetical prote
409	29.5	50.0	343	2	T02071	glyceraldehyde-3-p	482	29	49.2	280	2	T27758	hypothetical prote
410	29.5	50.0	373	2	T34126	hypothetical prote	483	29	49.2	282	2	C95181	degv family protei
411	29.5	50.0	430	2	S50604	AST2 protein - yea	484	29	49.2	282	2	F98048	conserved hypotet
412	29.5	50.0	442	2	S50062	cell wall glycopro	485	29	49.2	284	2	G96660	protein F2k11.8 [i
413	29.5	50.0	1367	2	H82874	conserved hypotet	486	29	49.2	287	2	T25064	hypothetical prote
414	29.5	50.0	7962	2	I38346	elastic titin - hu	487	29	49.2	287	2	C97546	hypothetical prote
415	29	49.2	30	2	PQ0444	hypothetical prote	488	29	49.2	287	2	AE2765	cytochrome-c oxida
416	29	49.2	57	2	D81152	hypothetical prote	489	29	49.2	288	2	G69387	conserved hypotet
417	29	49.2	60	2	A54369	fibrinogen recepto	490	29	49.2	292	2	A97034	transcription regu
418	29	49.2	83	2	AE1851	hypothetical prote	491	29	49.2	296	2	T14672	hypothetical prote
419	29	49.2	86	2	A13131	hypothetical prote	492	29	49.2	296	2	T36437	probable ABC-type
420	29	49.2	110	2	E85547	probable transcrip	493	29	49.2	297	1	S49348	cytochrome-c oxida
421	29	49.2	117	2	S28161	nerve growth facto	494	29	49.2	297	2	T27584	hypothetical prote
422	29	49.2	120	2	T07765	disease resistance	495	29	49.2	299	2	A10188	ATP phosphoribosyl
423	29	49.2	124	2	S24696	Ig heavy chain V6	496	29	49.2	299	2	T14943	hypothetical prote
424	29	49.2	125	2	A26312	nerve growth facto	497	29	49.2	300	2	AF3092	conserved hypotet
425	29	49.2	132	2	AB1757	hypothetical prote	498	29	49.2	300	2	D98194	hypothetical prote
426	29	49.2	134	2	AF3482	hypothetical prote	499	29	49.2	301	2	T27585	hypothetical prote
427	29	49.2	135	2	F64779	probable transcrip	500	29	49.2	302	2	A96661	unknown protein, 8
428	29	49.2	135	2	A90697	probable transcrip							
429	29	49.2	138	2	AF0564	copper efflux regu							
430	29	49.2	142	2	A47142	cystatin D precurs							
431	29	49.2	142	2	D84300	bacterio-opsin lin							
432	29	49.2	145	2	A55932	galectin-5 - rat							
433	29	49.2	145	2	AD1215	hypothetical prote							
434	29	49.2	145	2	AH1568	hypothetical prote							
435	29	49.2	149	2	S12913	cystatin - fruit f							
436	29	49.2	151	2	A05033	rpcC protein homol							
437	29	49.2	157	1	EDXF3	immediate-early pr							
438	29	49.2	161	2	T10853	V4ir protein - Rhi							
439	29	49.2	164	2	S63995	peptidylprolyl iso							
440	29	49.2	165	2	B38388	peptidylprolyl iso							
441	29	49.2	171	2	T07070	peptidylprolyl iso							
442	29	49.2	172	2	T07950	peptidylprolyl iso							
443	29	49.2	172	2	T06073	peptidylprolyl iso							
444	29	49.2	172	2	S54833	peptidylprolyl iso							
445	29	49.2	173	2	H69079	conserved hypotet							
446	29	49.2	173	2	AD0718	probable lipoprote							
447	29	49.2	174	2	E84597	cyclophilin (CYP2)							
448	29	49.2	174	2	T07072	peptidylprolyl iso							
449	29	49.2	174	2	JE0323	gammaS-crystallin							
450	29	49.2	178	2	A93279	invasion protein A							
451	29	49.2	179	2	S48018	peptidylprolyl iso							
452	29	49.2	182	2	I67436	interleukin-1-beta							
453	29	49.2	191	2	A71814	hypothetical prote							
454	29	49.2	200	2	T42678	hypothetical prote							
455	29	49.2	201	2	A30833	ribosomal protein							
456	29	49.2	205	2	B97584	30S ribosomal prot							
457	29	49.2	205	2	AB2805	30S ribosomal prot							
458	29	49.2	205	2	F87560	ribosomal protein							
459	29	49.2	208	2	E95338	hypothetical prote							
460	29	49.2	213	2	S43723	lactase (EC 3.2.1.							
461	29	49.2	213	2	S43722	lactase (EC 3.2.1.							
462	29	49.2	214	2	S65052	peptidyl-specific pr							
463	29	49.2	216	2	E84450	hypothetical prote							
464	29	49.2	218	2	T08737	hypothetical prote							
465	29	49.2	227	2	T47186	hypothetical prote							
466	29	49.2	229	2	T09215	H+-exporting ATPas							
467	29	49.2	235	2	S14481	nerve growth facto							

RESULT 1

GB3840

ferric uptake regulation protein BH1527 [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004

C:Accession: GB3840

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hiran

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and c

A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: GB3840

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-150 <STO>

A:Cross-references: UNIPROT:Q9KCP2; GB:AP001512; GB:BA000004; NID:g10174030; PIDN:BA0524

A:Experimental source: strain C-125

C:Genetics:

C:Superfamily: ferric uptake regulator

Query Match 67.8%; Score 40; DB 2; Length 150;

Best Local Similarity 66.7%; Pred. No. 4.3;

Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDSK 10

Db 141 HRCQDESK 149

||||:||||

RESULT 2

A39130

nicotinate phosphoribosyltransferase (EC 2.4.2.11) - Salmonella typhimurium

C:Species: Salmonella typhimurium

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C:Accession: A39130

R;Vinitsky, A.; Teng, H.; Grubmeyer, C.T.

J. Bacteriol. 173, 536-540, 1991

A;Title: Cloning and nucleic acid sequence of the Salmonella typhimurium pncB gene and a reference sequence

A;Reference number: A39130; MUID:91100340; PMID:1987148

A;Accession: A39130

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-400 <VIN>

A;Cross-references: UNIPROT:P22253; GB:M55986; NID:g154268; PIDN:AAA27190.1; PID:g154269

C;Genetics:

A;Gene: pncB

C;Function:

A;Pathway: nicotinate and nicotinamide metabolism

C;Superfamily: nicotinate phosphoribosyltransferase

C;Keywords: glycosyltransferase; pentosyltransferase

Query Match 67.8%; Score 40; DB 1; Length 400;

Best Local Similarity 85.7%; Pred. No. 11;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDD 8

Db 39 FRCRGDD 45

RESULT 3

JQ0756

nicotinate phosphoribosyltransferase (EC 2.4.2.11) - Escherichia coli (strain K-12)

C;Species: Escherichia coli

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C;Accession: JQ0756; B64833

R;Wubolts, M.G.; Terpstra, P.; van Beilen, J.B.; Kingma, J.; Meesters, H.A.R.; Witholt, J. Biol. Chem. 265, 17665-17672, 1990

A;Title: Variation of cofactor levels in Escherichia coli; sequence analysis and expression

A;Reference number: JQ0756; MUID:91009224; PMID:2211655

A;Accession: JQ0756

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-400 <WUB>

A;Cross-references: UNIPROT:P18133; GB:J05568; NID:g147306; PIDN:AAA24400.1; PID:g147307

A;Experimental source: strain GEC70

R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A;Title: The complete genome sequence of Escherichia coli K-12.

A;Reference number: A64720; MUID:97426617; PMID:9278503

A;Accession: B64833

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-400 <BLAT>

A;Cross-references: GB:AE000195; GB:U00096; NID:g1787156; PIDN:AAC74017.1; PID:g1787162;

A;Experimental source: strain K-12, substrain MG1655

C;Genetics:

A;Gene: pncB

C;Function:

A;Pathway: nicotinate and nicotinamide metabolism

C;Superfamily: nicotinate phosphoribosyltransferase

C;Keywords: glycosyltransferase; pentosyltransferase

Query Match 67.8%; Score 40; DB 1; Length 400;

Best Local Similarity 85.7%; Pred. No. 11;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDD 8

Db 39 FRCRGDD 45

RESULT 4

AB0617

nicotinate phosphoribosyltransferase [imported] - Salmonella enterica subsp. enterica serovar Typhimurium

C;Species: Salmonella enterica subsp. enterica serovar Typhi

A;Note: this species has also been called Salmonella typhi

C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002

C;Accession: AB0617

R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Cronin, P.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar

A;Reference number: AB0502; MUID:21534947; PMID:11677608

A;Accession: AB0617

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-400 <PAR>

A;Cross-references: GB:AL513382; PIDN:CAD05404.1; PID:g16502165; GSPDB:GN00176

C;Genetics:

A;Gene: STY1010

C;Superfamily: nicotinate phosphoribosyltransferase

Query Match 67.8%; Score 40; DB 2; Length 400;

Best Local Similarity 85.7%; Pred. No. 11;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDD 8

Db 39 FRCRGDD 45

RESULT 5

F90755

nicotinate phosphoribosyltransferase [imported] - Escherichia coli (strain O157:H7, subsp

C;Species: Escherichia coli

C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004

C;Accession: F90755

R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ighii, K.; Yokoyama, K.; Han, C.G.;

gaawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom

A;Reference number: A99629; MUID:21156231; PMID:11258796

A;Accession: F90755

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-400 <HAY>

A;Cross-references: UNIPROT:Q8XDE8; GB:BA000007; PIDN:BAE34437.1; PID:g13360473; GSPDB:GP

A;Experimental source: strain O157:H7, substrain RIMD 0509952

C;Genetics:

A;Gene: ECs1014

C;Superfamily: nicotinate phosphoribosyltransferase

Query Match 67.8%; Score 40; DB 2; Length 400;

Best Local Similarity 85.7%; Pred. No. 11;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDD 8

Db 39 FRCRGDD 45

RESULT 6

D85619

nicotinate phosphoribosyltransferase [imported] - Escherichia coli (strain O157:H7, subsp

C;Species: Escherichia coli

C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004

C;Accession: D85619

R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, A.; Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, Nature 409, 529-533, 2001

A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A;Reference number: AB5480; MUID:21074935; PMID:11206551

A;Accession: D85619

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-400 <STO>

A;Cross-references: UNIPROT:Q8XDE8; GB:AE005174; NID:g12514106; PIDN:AGS55416.1; GSPDB:GP

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: pncB

C:Superfamily: nicotinate phosphoribosyltransferase

Query Match 67.8%; Score 40; DB 2; Length 400;
Best Local Similarity 85.7%; Pred. No. 11;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDD 8
:|||||
Db 39 FRCRGDD 45

RESULT 7

A70669 probable acid-CoA ligase (EC 6.2.1.1.-) fadD22 [similarity] - Mycobacterium tuberculosis

N:Alternate names: hypothetical protein pks002a

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C:Accession: A70669; S73073

R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: A70669

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-705 <COL>

A:Cross-references: UNIPROT:P96283; GB:Z83858; GB:AL123456; NID:G3261675; PIDN:CAB06101.

A:Experimental source: strain H37Rv

R:Smith, D.R.; Robison, K.

submitted to the EMBL Data Library, September 1994

A:Description: Mycobacterium tuberculosis cosmid tbc2.

A:Reference number: S73053

A:Accession: S73073

A:Molecule type: DNA

A:Residues: 1-486, 'QQ', 489-705 <SMI>

A:Cross-references: EMBL:U00024; NID:G560506; PIDN:AAA50930.1; PID:G560509

C:Genetics:

A:Gene: fadD22

C:Keywords: acid-thiol ligase; carrier protein; phosphopantetheine; phosphoprotein

F:51-486/Domain: acetate-CoA ligase homology <ACI>

F:544-615/Domain: acyl carrier protein homology <ACPI>

F:579/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 67.8%; Score 40; DB 2; Length 705;
Best Local Similarity 66.7%; Pred. No. 18;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 YRCRGDDSK 10
|||||
Db 383 YRCRADDTTE 391

RESULT 8

hypothetical protein F52H3.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T22521

R:Gardner, A.

submitted to the EMBL Data Library, October 1995

A:Reference number: Z19575

A:Accession: T22521

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-179 <WIL>

A:Cross-references: UNIPROT:Q20683; EMBL:Z66512; PIDN:CAA91325.1; GSPDB:GN000020; CESP:F5

A:Experimental source: clone F52H3

C:Genetics:

A:Gene: CESP:F52H3.5

A:Map position: 2

A:Introns: 68/2; 95/1; 135/3

Query Match 66.1%; Score 39; DB 2; Length 179;
Best Local Similarity 77.8%; Pred. No. 7.8;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 YRCRGDDSK 10
|||||
Db 125 YRLRGDDDK 133

RESULT 9

T27085

hypothetical protein Y51A2D.17 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 16-Aug-2004

C:Accession: T27085; T27320

R:McMurray, A.

submitted to the EMBL Data Library, January 1998

A:Reference number: Z20307

A:Accession: T27085

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-389 <WIL>

A:Cross-references: UNIPROT:Q9XTG0; UNIPROT:Q9GTD6; EMBL:AL021497; PIDN:CAA16413.1; GSPDI

A:Experimental source: clone Y51A2D

R:McMurray, A.

submitted to the EMBL Data Library, August 1997

A:Reference number: Z20343

A:Accession: T27320

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-389 <WIL>

A:Cross-references: EMBL:Z98877; PIDN:CAB11571.1; GSPDB:GN00023; CESP:Y51A2D.17

A:Experimental source: clone Y69H2

C:Genetics:

A:Gene: CESP:Y51A2D.17

A:Map position: 5

A:Introns: 29/2; 58/1; 80/1; 162/3; 204/3; 228/1; 305/1; 356/3

C:Superfamily: erba transforming protein homology

Query Match 64.4%; Score 38; DB 2; Length 389;

Best Local Similarity 100.0%; Pred. No. 24;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGD 7
|||||
Db 45 YRCRGD 50

RESULT 10

H82528

L-ascorbate oxidase XP2677 [imported] - Xylella fastidiosa (strain 9a5c)

C:Species: Xylella fastidiosa

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C:Accession: H82528

R:anonymus, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequenc

Nature 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: AB2515; MUID:20365717; PMID:10910347

A:Note: for a complete list of authors see reference number A59328 below

A:Accession: H82528

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-721 <SIM>

A:Cross-references: UNIPROT:Q9PA43; GB:AE003849; NID:G9107904; PIDN:AAF85474

A:Experimental source: strain 9a5c

R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; AJ

Briones, M.R.S.; Bueno, M.R.F.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carreir, H.

as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranae, E.B.; Laigh Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Mizacca, E.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawaak A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira M.; Tuhakko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z A:Content: annotation
A:Reference number: A59328
C:Genetics:
A:Gene: XP2677

Query Match 64.4%; Score 38; DB 2; Length 721;
Best Local Similarity 66.7%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 YRCRGDSK 10
|||:|
Db 282 YACRGNDGK 290

RESULT 11
MWXR31
lambda 3 protein - reovirus type 1 (strain Lang)
N:Alternate names: minor core protein
C:Species: reovirus type 1
C>Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
C:Accession: A30121
R:Wiener, J.R.; Joklik, W.K.
Virology 169, 194-203, 1989
A:Title: The sequences of the reovirus serotype 1, 2, and 3 L1 genome segments and analysis
A:Reference number: A94390; MUID:89163254; PMID:2922925
A:Accession: A30121
A:Molecule type: Genomic RNA
A:Residues: 1-1267 <WIE>
A:Cross-references: UNIPROT:P17376; GB:M24734; NID:G499863
A:Note: This sequence, which matches the sequence attributed to type 1 in Fig. 2, matches the translations in entries REO1LAM3P and REO3LAM3P now differ only by the sequence correction
C:Comment: See also PIR:MWXR31.
C:Genetics:
A:Map position: segment L1
C:Superfamily: reovirus lambda 3 protein
C:Keywords: core protein

Query Match 64.4%; Score 38; DB 1; Length 1267;
Best Local Similarity 75.0%; Pred. No. 71;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDD 8
|||:|
Db 728 NYVCGDD 735

RESULT 12
MWXR32
lambda 3 protein - reovirus type 2 (strain D5/Jones)
N:Alternate names: minor core protein
C:Species: reovirus type 2
A:Note: host Homo sapiens (man)
C>Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
C:Accession: B30121
R:Wiener, J.R.; Joklik, W.K.
Virology 169, 194-203, 1989
A:Title: The sequences of the reovirus serotype 1, 2, and 3 L1 genome segments and analysis
A:Reference number: A94390; MUID:89163254; PMID:2922925
A:Accession: B30121
A:Molecule type: Genomic RNA
A:Residues: 1-1267 <WIE>
A:Cross-references: UNIPROT:P17377; GB:M31057; NID:G499865; PIDN:AAA47245.1; PID:G499866
C:Genetics:
A:Map position: segment L1

C:Superfamily: reovirus lambda 3 protein
C:Keywords: core protein

Query Match 64.4%; Score 38; DB 1; Length 1267;
Best Local Similarity 75.0%; Pred. No. 71;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDD 8
|||:|
Db 728 NYVCGDD 735

RESULT 13
MWXR33
lambda 3 protein - reovirus type 3 (strain Dearing)
N:Alternate names: minor core protein
C:Species: reovirus type 3
C>Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
C:Accession: C30121
R:Wiener, J.R.; Joklik, W.K.
Virology 169, 194-203, 1989
A:Title: The sequences of the reovirus serotype 1, 2, and 3 L1 genome segments and analysis
A:Reference number: A94390; MUID:89163254; PMID:2922925
A:Accession: C30121
A:Molecule type: Genomic RNA
A:Residues: 1-1267 <WIE>
A:Cross-references: UNIPROT:P17378; GB:M31058; NID:G499867; GB:M24734; NID:G499863; PIDN: nrides REO1LAM3P and REO3LAM3P now differ only by the sequence correction apparently made
C:Comment: See also PIR:MWXR31.
C:Genetics:
A:Map position: segment L1
C:Superfamily: reovirus lambda 3 protein
C:Keywords: core protein

Query Match 64.4%; Score 38; DB 1; Length 1267;
Best Local Similarity 75.0%; Pred. No. 71;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDD 8
|||:|
Db 728 NYVCGDD 735

RESULT 14
S75313
serine-tRNA ligase (EC 6.1.1.11) - Synecocystis sp. (strain PCC 6803)
N:Alternate names: protein slr1703; seryl-tRNA synthetase
C:Species: Synecocystis sp.
A:Variety: PCC 6803
C>Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S75313
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O.K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S75313
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-430 <KAN>
A:Cross-references: UNIPROT:P73201; EMBL:D90904; GB:AB001339; NID:G1652225; PIDN:BAAL722;
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
A:Gene: sers
A:Start codon: GTG
C:Function:
A:Description: charges tRNA(Ser) with serine
A:Pathway: protein biosynthesis
C:Superfamily: serine-tRNA ligase
C:Keywords: aminoacyl-tRNA synthetase; ATP; ligase; protein biosynthesis

```

Query Match          62.7%; Score 37; DB 2; Length 430;
Best Local Similarity 62.5%; Pred. No. 40;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYRCRGDD 8
DB 225 SFQCRGDD 232

RESULT 15
C64398
hypothetical protein homolog MJ0787 - Methanococcus jannaschii
C:Species: Methanococcus jannaschii
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C:Accession: C64398
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A:Reference number: A64300; MUID:96237999; PMID:8688087
A:Accession: C64398
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-504 <BUL>
A:Cross-references: UNIPROT:Q58197; GB:U67523; GB:L77117; NID:g2826319; PIDN:AAB98783.1;
C:Genetics:
A:Map position: FOR710775-712289

Query Match          62.7%; Score 37; DB 2; Length 504;
Best Local Similarity 66.7%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 YRCRGDDSK 10
DB 463 YPCNGDDKK 471

RESULT 16
T31420
C-terminal domain-binding protein rA8 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T31420
R:Yuryev, A.; Patturajan, M.; Liington, Y.; Joshi, R.V.; Gentile, C.; Gebara, M.; Cord
Proc. Natl. Acad. Sci. U.S.A. 93, 6975-6980, 1996
A:Title: The C-terminal domain of the largest subunit of RNA polymerase II interacts wit
A:Reference number: Z21024; MUID:96293459; PMID:8692929
A:Accession: T31420
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1268 <YUR>
A:Cross-references: UNIPROT:Q63623; EMBL:U49055; NID:g1438529; PID:g1438530; PIDN:AA526
A:Experimental source: hippocampus

Query Match          62.7%; Score 37; DB 2; Length 1268;
Best Local Similarity 85.7%; Pred. No. 11e-02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 YRCRGDD 8
DB 100 YRCPGDD 106

RESULT 17
T35013
probable membrane protein - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C:Accession: T35013
R:Seeger, S.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, June 1999

A:Reference number: Z21565
A:Accession: T35013
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-347 <SEE>
A:Cross-references: UNIPROT:Q9XAN5; EMBL:AL079355; PIDN:CAB45569.1; GSPDB:GN000070; SCODEI
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCORDB:SC4C6.05C
C:Superfamily: Streptomyces coelicolor probable membrane protein SC4C6.05C

Query Match          61.0%; Score 36; DB 2; Length 347;
Best Local Similarity 66.7%; Pred. No. 51;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 YRCRGDDSK 10
DB 206 YKCPGDTSK 214

RESULT 18
E93360
4-hydroxyphenylacetate 3-monooxygenase (EC 1.14.13.3) hpaA-2 - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: E93360
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson,
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.N
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeor
A:Reference number: A69250; MUID:98049343; PMID:9389475
A:Accession: E93360
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-461 <KLE>
A:Cross-references: UNIPROT:O29377; GB:AE001043; GB:AE000782; NID:g2689366; PIDN:AAB9035
C:Superfamily: Escherichia coli 4-hydroxyphenylacetate 3-monooxygenase large chain
C:Keywords: FAD; NAD; oxidoreductase

Query Match          61.0%; Score 36; DB 1; Length 461;
Best Local Similarity 66.7%; Pred. No. 66;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYRCRGDD 9
DB 94 NYRCTGCDA 102

RESULT 19
E72376
hypothetical protein TM0434 - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: E72376; F72298
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey,
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; MUID:99287316; PMID:10360571
A:Accession: E72376
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-470 <ARN>
A:Cross-references: UNIPROT:Q9WYR5; GB:AE001722; GB:AE000512; NID:g4980938; PIDN:AAD3551
A:Experimental source: strain MSB8
A:Accession: F72298
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-415, 'A', 417-464, 'GS' <AR2>

```

A;Cross-references: GB:AE001767; GB:AE000512; NID:g4981611; PIDN:AAD36145.1; PID:g4981611
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM0434, TM1068
C;Superfamily: melibiose-specific alpha-galactosidase

Query Match 61.0%; Score 36; DB 2; Length 470;
Best Local Similarity 71.4%; Pred. No. 67;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDD 8
:|||||
Db 214 FRCRGED 220

RESULT 20
AD3382
ABC transporter-associated protein BME11042 [imported] - Brucella melitensis (strain 16M
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C;Accession: AD3382
R;BelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujaer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goldsman, E.; Salkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A;Reference number: AD3252; PMID:11756688
A;Accession: AD3382
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-507 <KUR>
A;Cross-references: UNIPROT:Q8YGM3; GB:AE008917; PIDN:AAL52223.1; PID:g17983006; GSPDB:G
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BME11042
A;Map position: I
C;Superfamily: conserved hypothetical protein b1683

Query Match 61.0%; Score 36; DB 2; Length 507;
Best Local Similarity 85.7%; Pred. No. 72;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 CRGDDSK 10
|||||
Db 319 CRGDSK 325

RESULT 21
I49350
breast/ovarian cancer susceptibility protein BRCA1 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-May-2004
C;Accession: I49350
R;Marquis, S.T.; Rajan, J.V.; Wynshaw-Boris, A.; Xu, J.; Yin, G.Y.; Abel, K.J.; Weber, E
Nature Genet. 11, 17-26, 1995
A;Title: The developmental pattern of Brcal expression implies a role in differentiation
A;Reference number: I49350; MUID:96021028; PMID:7550308
A;Accession: I49350
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1812 <RES>
A;Cross-references: EMBL:U31625; NID:g988213; PIDN:AAB17114.1; PID:g988214
C;Genetics:
A;Gene: Brcal
C;Superfamily: transcriptional regulator, BRCA1 type; RING finger homology
C;Keywords: zinc
F;20-70/Domain: RING finger homology <RNG>

Query Match 61.0%; Score 36; DB 2; Length 1812;
Best Local Similarity 66.7%; Pred. No. 2.3e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDS 9
:|||||

Db 1786 SYRCRGLDA 1794

RESULT 22
C90460
hypothetical protein SSO2828 [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C;Accession: C90460
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-y
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: C90460
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-138 <KUR>
A;Cross-references: UNIPROT:Q97V09; GB:AE006641; NID:gl3816183; PIDN:AAK42938.1; GSPDB:GN
C;Genetics:
A;Gene: SSO2828

Query Match 59.3%; Score 35; DB 2; Length 138;
Best Local Similarity 66.7%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 YRCRGDDSK 10
|||||
Db 45 YRCGDIGK 53

RESULT 23
AI0520
AmpD protein (anhydro-N-acetylmuramyl-tripeptide amidase) [imported] - Salmonella enteric
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Dec-2002
C;Accession: AI0520
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
. S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AI0520
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-187 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD01305.1; PID:g16501433; GSPDB:GN00176
C;Genetics:
A;Gene: STY0168
C;Superfamily: AmpD protein

Query Match 59.3%; Score 35; DB 2; Length 187;
Best Local Similarity 60.0%; Pred. No. 44;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
:|||||
Db 18 HYDCRPDDEK 27

RESULT 24
A46506
N leukocyte activation antigen M6 - human
N;Alternate names: basigin; collagenase stimulatory factor, tumor-derived
C;Species: Homo sapiens (man)
C;Date: 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
A;Accession: A46506; S13906; I52729; I55194; S19203
R;Kasinrerk, W.; Fiebigler, E.; Stefanova, I.; Baumruker, T.; Knapp, W.; Stockinger, H.
J. Immunol. 149, 847-854, 1992

A:Title: Human leukocyte activation antigen M6, a member of the Ig superfamily, is the s
A:Reference number: A46506; MUID:92340888; PMID:1634773
A:Accession: A46506
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-269 <RAS>
A:Cross-references: UNIPROT:P35613; EMBL:X64364; NID:G34448; PIDN:CAA45716.1; PID:G34449
A>Note: sequence extracted from NCBI backbone (NCBIP:109314)
R:Nabeshima, K.; Lane, W.S.; Biswas, C.
Arch. Biochem. Biophys. 285, 90-96, 1991
A:Title: Partial sequencing and characterization of the tumor cell-derived collagenase s
A:Reference number: S13906; MUID:91119430; PMID:1846736
A:Accession: S13906
A>Status: preliminary
A:Molecule type: protein
A:Residues: 22-36 <NAB>
R:Biswas, C.; Zhang, Y.; DeCastro, R.; Guo, H.; Nakamura, T.; Kataoka, H.; Nabeshima, K.
Cancer Res. 55, 434-439, 1995
A:Title: The human tumor cell-derived collagenase stimulatory factor (renamed EMMPRIN) i
A:Reference number: I52729; MUID:95112304; PMID:7812975
A:Accession: I52729
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-269 <RES>
A:Cross-references: GB:I10240; NID:9409356; PIDN:AAA68936.1; PID:G409357
R:Miyauchi, T.; Masuzawa, Y.; Muramatsu, T.
J. Biochem. 110, 770-774, 1991
A:Title: The basigin group of the immunoglobulin superfamily: complete conservation of a
A:Reference number: I55194
A:Accession: I55194
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-269 <RE2>
A:Cross-references: GB:D45131; NID:gl304103; PIDN:BAA08109.1; PID:G633069
C:Genetics:
A:Gene: GDB:BSG
A:Cross-references: GDB:228978; OMIM:109480
A:Map position: 19p13.3-19p13.3
C:Superfamily: cell surface glycoprotein HT7; immunoglobulin homology
C:Keywords: transmembrane protein

Query Match 59.3%; Score 35; DB 2; Length 269;
Best Local Similarity 66.7%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 YRCRGDSK 10
||| |||
DB 183 YRCNGTSSK 191

RESULT 25
T32736
hypothetical protein F54D7.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T32736
R:Dante, M.; Twyman, B.
submitted to the EMBL Data Library, December 1997
A:Description: The sequence of C. elegans cosmid F54D7.
A:Reference number: Z21218
A:Accession: T32736
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-274 <DAN>
A:Cross-references: EMBL:AF039712; PIDN:AAB96718.1; GSPDB:GN00019; CESP:F54D7.4
A:Experimental source: strain Bristol N2; clone F54D7
C:Genetics:
A:Gene: CESP:F54D7.4
A:Map position: 1
A:Introns: 10/1; 54/2; 78/3; 124/1; 203/1; 258/2

Query Match 59.3%; Score 35; DB 2; Length 274;
Best Local Similarity 75.0%; Pred. No. 62;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 YRCRGDS 9
||| |||
DB 228 YRCIGDNS 235

RESULT 26
T20546
hypothetical protein F07C6.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T20546
R:Steward, C.
submitted to the EMBL Data Library, February 1996
A:Reference number: Z19290
A:Accession: T20546
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-329 <WIL>
A:Cross-references: UNIPROT:O19157; EMBL:Z69659; PIDN:CAA93483.1; GSPDB:GN00022; CESP:F07C6
A:Experimental source: clone F07C6
C:Genetics:
A:Gene: CESP:F07C6.1
A:Map position: 4
A:Introns: 14/2; 47/2; 85/1; 146/1; 226/3; 299/2

Query Match 59.3%; Score 35; DB 2; Length 329;
Best Local Similarity 85.7%; Pred. No. 74;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 YRCRGDD 8
||| |||
DB 287 YRCRGCD 293

RESULT 27
T24127
probable chitinase (EC 3.2.1.14) precursor R10D12.15 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T24127; T25316; T25318
R:Percy, C.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19842
A:Accession: T24127
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-453 <WIL>
A:Cross-references: UNIPROT:P92013; EMBL:Z81109; PIDN:CAB03255.1; GSPDB:GN00023; CESP:R10D12
A:Experimental source: clone R10D12
R:Percy, C.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z20014
A:Accession: T25316
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-453 <W12>
A:Cross-references: EMBL:Z82054; PIDN:CAB04840.1; GSPDB:GN00023; CESP:T26F2.1
A:Experimental source: clone T26F2
A:Accession: T25318
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-453 <W13>
A:Cross-references: EMBL:Z82054; PIDN:CAB04842.1; GSPDB:GN00023; CESP:R10D12.15
A:Experimental source: clone T26F2
C:Genetics:
A:Gene: CESP:R10D12.15; CESP:T26F2.1
A:Map position: 5
A:Introns: 23/1; 118/3; 150/2; 237/3; 369/1
C:Keywords: glycosidase; hydrolase

Query Match 59.3%; Score 35; DB 2; Length 453;

Best Local Similarity 60.0%; Pred. No. 99;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
| | | | |
Db 325 NNECSGSDSK 334

RESULT 28
B87484
anthranilate synthase component I [imported] - Caulobacter crescentus
C/Species: Caulobacter crescentus
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C/Accession: B87484
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A/Title: Complete Genome Sequence of Caulobacter crescentus.
A/Reference number: A87249; MUID:21173698; PMID:11259647
A/Accession: B87484
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-513 <STO>
A/Cross-references: UNIPROT:Q9A731; GB:AE005673; NID:gl3423344; PIDN:AAK23870.1; GSPDB:C
C/Genetics:
A/Gene: CCI895
C/Superfamily: anthranilate synthase component I

Query Match 59.3%; Score 35; DB 2; Length 513;
Best Local Similarity 55.6%; Pred. No. 1.1e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 YRCRGDDSK 10
: | | | | :
Db 73 WRCRGDQAE 81

RESULT 29
T10664
serine/threonine-specific protein kinase homolog F6E21.20 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C/Accession: T10664
R;Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancro
submitted to the Protein Sequence Database, June 1999
A/Reference number: Z16533
A/Accession: T10664
A/Molecule type: DNA
A/Residues: 1-656 <BEV>
A/Cross-references: UNIPROT:Q9M092; EMBL:AL049914; GSPDB:GN000062; ATSP:F6E21.20
A/Experimental source: cultivar Columbia; BAC clone F6E21
C/Genetics:
A/Gene: ATSP:F6E21.20
A/Map position: 4
A/Introns: 206/1; 243/1
C/Superfamily: Arabidopsis probable serine/threonine-specific protein kinase PRO25; prob

Query Match 59.3%; Score 35; DB 2; Length 656;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 YRCRGDDSK 10
| | | | |
Db 230 YRCVGDKTK 238

RESULT 30
I37892
IL12 receptor component - human
C/Species: Homo sapiens (man)
C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C/Accession: I37892

R;Chua, A.O.; Chizzonite, R.; Desai, B.B.; Truitt, T.P.; Nunes, P.; Minetti, L.J.; Warrie
J. Immunol. 153, 128-136, 1994
A/Title: Expression cloning of a human IL-12 receptor component. A new member of the cyto
A/Reference number: I37892; MUID:94267217; PMID:7911493
A/Accession: I37892
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-662 <RES>
A/Cross-references: UNIPROT:P42701; EMBL:U03187; NID:gs07150; PIDN:AAA21340.1; PID:gs0715

Query Match 59.3%; Score 35; DB 2; Length 662;
Best Local Similarity 75.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 RCRGDDSK 10
| | | | |
Db 484 RCRDEDSK 491

RESULT 31
T37533
coronin-like protein - fission yeast (Schizosaccharomyces pombe)
C/Species: Schizosaccharomyces pombe
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C/Accession: T37533
R;Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, August 1997
A/Reference number: Z21721
A/Accession: T37533
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1323 <MUR>
A/Cross-references: UNIPROT:O13686; EMBL:Z98595; PIDN:CAB1184.1; GSPDB:GN000066; SPDB:SP
A/Experimental source: strain 972h; cosmid c11E3
C/Genetics:
A/Gene: SPDB:SPAC11E3.05
A/Map position: 1

Query Match 59.3%; Score 35; DB 2; Length 1323;
Best Local Similarity 71.4%; Pred. No. 2.6e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDD 8
: | | | | :
Db 512 WRCRGD 518

RESULT 32
B81364
probable periplasmic protein Cj0906c [imported] - Campylobacter jejuni (strain NCTC 1116f
C/Species: Campylobacter jejuni
C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C/Accession: B81364
R;Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell
Nature 403, 665-668, 2000
A/Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
A/Reference number: A81250; MUID:20150912; PMID:10688204
A/Accession: B81364
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-325 <PAR>
A/Cross-references: UNIPROT:Q9PP25; GB:AL139076; GB:AL111168; NID:g6968128; PIDN:CAB73164
A/Experimental source: serotype O2, strain NCTC 11168
C/Genetics:
A/Gene: Cj0906c
C/Superfamily: Helicobacter pylori hypothetical protein jhp0467

Query Match 58.5%; Score 34.5; DB 2; Length 325;
Best Local Similarity 52.9%; Pred. No. 90;
Matches 9; Conservative 0; Mismatches 1; Indels 7; Gaps 1;

QY 1 NYRCRGD-----DSK 10

```

Db      300 NYRFRGDKILYVKIDSK 316
|||||
RESULT 33
H70232
hypothetical protein BBG16 - Lyme disease spirochete plasmid G/lp28-2
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004
C:Accession: H70232
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White-
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: H70232
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-112 <KLE>
A:Cross-references: UNIPROT:O50739; GB:AE000786; NID:g2690008; PIDN:AAC66064.1; PID:g269
A:Experimental source: strain B31
C:Genetics:
C:Superfamily: Lyme disease spirochete plasmid G/lp28-2 hypothetical protein BBG16

Query Match      57.6%; Score 34; DB 2; Length 112;
Best Local Similarity 50.0%; Pred. No. 42;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 NYRCRGDDSK 10
      |||||
Db      77 NIRCQGDQKE 86
|||||

RESULT 34
T04118
mitochondrial processing peptidase homolog - rice (fragment)
C:Species: Oryza sativa (rice)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T04118
R:Uchiyama, H.
submitted to the EMBL Data Library, November 1993
A:Reference number: Z14944
A:Accession: T04118
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-122 <UCH>
A:Cross-references: UNIPROT:Q40738; EMBL:D25241; PIDN:BAA04964.1
C:Genetics:
A:Gene: SS656
C:Superfamily: mitochondrial processing peptidase alpha chain
C:Keywords: mitochondrion

Query Match      57.6%; Score 34; DB 2; Length 122;
Best Local Similarity 50.0%; Pred. No. 46;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 NYRCRGDDSK 10
      |||||
Db      48 DYRCQADSK 57
|||||

RESULT 35
F81439
translation initiation factor IF-3 Cj0207 [imported] - Campylobacter jejuni (strain NCTC
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: F81439
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kerley, J.M.; Churcher, C.; Baeham, D.; Chillin
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrel
Nature 403, 665-668, 2000

```

```

A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hype
A:Reference number: A81250; MUID:20150912; PMID:10688204
A:Accession: F81439
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-172 <PAR>
A:Cross-references: UNIPROT:Q9PIS2; GB:ALJ39074; GB:ALJ11168; NID:g6967505; PIDN:CAB7269
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: infC; Cj0207
C:Superfamily: translation initiation factor IF-3

Query Match      57.6%; Score 34; DB 2; Length 172;
Best Local Similarity 75.0%; Pred. No. 62;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      3 RCRGDDSK 10
      |||||
Db      19 RCVGDDGK 26
|||||

RESULT 36
AC3270
hypothetical exported protein BMEI0144 [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C:Accession: AC3270
R:DelVecchio, V.G.; Kaputral, V.; Redkar, R.J.; Patra, G.; Mujter, C.; Loos, T.; Ivanova, I
; Mazur, M.; Goltzman, E.; Selkov, E.; Elzei, P.H.; Hagius, S.; O'Callaghan, D.; Letessc
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:11756688
A:Accession: AC3270
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-180 <KUR>
A:Cross-references: UNIPROT:Q8YJ2; UNIPROT:Q8FYG1; GB:AE008917; PIDN:AAL51326.1; PID:g1
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI0144
A:Map position: 1

Query Match      57.6%; Score 34; DB 2; Length 180;
Best Local Similarity 55.6%; Pred. No. 65;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 YRCRGDDSK 10
      |||||
Db      97 YNCKGDGSE 105
|||||

RESULT 37
G64320
hypothetical protein MJ0166 - Methanococcus jannaschii
C:Species: Methanococcus jannaschii
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C:Accession: G64320
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, J
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii.
A:Reference number: A64300; MUID:96337999; PMID:8688087
A:Accession: G64320
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-255 <BUJ>
A:Cross-references: UNIPROT:Q57630; GB:U67473; GB:L77117; NID:g2826256; PIDN:AAB98148.1;
C:Genetics:
A:Map position: REV170194-169427
C:Superfamily: Pyrococcus horikoshii hypothetical protein PH0006

```

```

Query Match          57.6%; Score 34; DB 2; Length 255;
Best Local Similarity 71.4%; Pred. No. 89;
Matches 5; Conservative 1; Mismatches 0; Gaps 0;

QY 4 CRGDDSK 10
    |||||
Db 233 CRGDDKR 239

RESULT 38
E84427
hypothetical protein At2g01660 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: E84427
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: E84427
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-259 <STO>
A:Cross-references: UNIPROT:Q9ZU94; GB:AE002093; NID:g4220482; PIDN:AAD12705.1; GSPDB:GN
C:Genetics:
A:Gene: At2g01660
A:Map position: 2

Query Match          57.6%; Score 34; DB 2; Length 259;
Best Local Similarity 83.3%; Pred. No. 91;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGD 7
    |:|||
Db 82 YQCRGD 87

RESULT 39
C96731
unknown protein F5A18.13 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: C96731
R:Theologas, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.E.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: C96731
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-270 <STO>
A:Cross-references: UNIPROT:Q9CAB9; GB:AE005173; NID:g6453891; PIDN:AAF09074.1; GSPDB:GN
C:Genetics:
A:Gene: F5A18.13
A:Map position: 1

Query Match          57.6%; Score 34; DB 2; Length 270;
Best Local Similarity 83.3%; Pred. No. 94;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGD 7
    |:|||
Db 87 YQCRGD 92

```

```

RESULT 40
T10610
hypothetical protein F21C20.30 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: T10610
R:Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Lemcke,
submitted to the Protein Sequence Database, June 1999
A:Reference number: Z16991
A:Accession: T10610
A:Molecule type: DNA
A:Residues: 1-273 <BEV>
A:Cross-references: UNIPROT:Q9SVH9; EMBL:AL080254; GSPDB:GN000662; ATSP:F21C20.30
A:Experimental source: cultivar Columbia; BAC clone F21C20
C:Genetics:
A:Gene: ATSP:F21C20.30
A:Map position: 4
A:Introns: 114/3

Query Match          57.6%; Score 34; DB 2; Length 273;
Best Local Similarity 83.3%; Pred. No. 95;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGD 7
    |:|||
Db 93 YQCRGD 98

RESULT 41
A95408
hypothetical protein Sma2147 [imported] - Sinorhizobium meliloti (strain 1021) magaplaemf
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C:Accession: A95408
R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bowse
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.;
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti
A:Reference number: A95262; MUID:21396509; PMID:11481432
A:Accession: A95408
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-296 <KUR>
A:Cross-references: UNIPROT:Q92XS2; GB:AE006469; PIDN:AAK65827.1; PID:g14524331; GSPDB:GN
A:Experimental source: strain 1021, megaplasmid pSymA
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, P.; Barloy-Hubler,
Pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.C.
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: Sma2147
A:Genome: plasmid

Query Match          57.6%; Score 34; DB 2; Length 296;
Best Local Similarity 50.0%; Pred. No. 1e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
    |:|||
Db 2 NEECKGDGGE 11

RESULT 42
T05437
hypothetical protein P7K2.10 - Arabidopsis thaliana (fragment)
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999

```

C:Accession: T05437
 R:Bevan, M.; Wedler, H.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Schueller submitted to the Protein Sequence Database, November 1998
 A:Reference number: Z15416
 A:Accession: T05437
 A:Molecule type: DNA
 A:Residues: 1-344 <BEV>
 A:Cross-references: EMBL:AL033545
 A:Experimental source: cultivar Columbia; BAC clone F7K2
 C:Genetics:
 A:Map position: 4
 A:Note: F7K2.10

Query Match 57.6%; Score 34; DB 2; Length 344;
 Best Local Similarity 60.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
 ||: ||: ||
 Db 267 NYKIFGDESK 276

RESULT 43
 H85256
 hypothetical protein AT4g22430 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
 C:Accession: H85256
 R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Nature 402, 769-777, 1999
 A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
 A:Reference number: A85001; MUID:20083488; PMID:10617198
 A:Accession: H85256
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-348 <STO>
 A:Cross-references: UNIPROT:Q9SUX6; GB:NC_001268; NID:g7269089; PIDN:CAB79198.1; GSPDB:G
 C:Genetics:
 A:Gene: AT4g22430
 A:Map position: 4

Query Match 57.6%; Score 34; DB 2; Length 348;
 Best Local Similarity 60.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
 ||: ||: ||
 Db 271 NYKIFGDESK 280

RESULT 44
 B96551
 hypothetical protein F11M15.18 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C:Accession: B96551
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, K.; Creasy, T.H.; Huijar, L.; ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C. A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzbarg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: B96551
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-375 <STO>
 A:Cross-references: UNIPROT:Q9SYD4; GB:AE005173; NID:g4836941; PIDN:AAD30643.1; GSPDB:GN
 C:Genetics:
 A:Gene: AT4g22430
 A:Map position: 4

A:Gene: F11M15.18
 A:Map position: 1
 C:Superfamily: Arabidopsis thaliana hypothetical protein F24M12.210

Query Match 57.6%; Score 34; DB 2; Length 375;
 Best Local Similarity 62.5%; Pred. No. 1.3e+02;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 YRCRGDDSK 9
 |: ||: ||
 Db 152 YKCEGDSS 159

RESULT 45
 B96189
 protein T25N20.9 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C:Accession: B96189
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, K.; Creasy, T.H.; Huijar, L.; ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C. A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzbarg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: B96189
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-393 <STO>
 A:Cross-references: UNIPROT:Q9ZVZ4; GB:AE005172; NID:g8778729; PIDN:AAF79737.1; GSPDB:GN
 C:Genetics:
 A:Gene: T25N20.9
 A:Map position: 1

Query Match 57.6%; Score 34; DB 2; Length 393;
 Best Local Similarity 75.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 RCRGDDSK 10
 ||: ||: ||
 Db 234 RCTGSDSK 241

RESULT 46
 C36786
 hypothetical protein 2 - ictalurid herpesvirus 1 (strain auburn 1)
 C:Species: ictalurid herpesvirus 1
 A:Note: host Ictalurus punctatus (channel catfish)
 C:Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 09-Jul-2004
 C:Accession: C36786
 R:Davidson, A.J.
 submitted to GenBank, January 1992
 A:Description: Channel catfish virus: a new type of herpesvirus.
 A:Reference number: A36804
 A:Accession: C36786
 A:Molecule type: DNA
 A:Residues: 1-407 <DAV>
 A:Cross-references: UNIPROT:Q00126; GB:M75136; NID:g331209; PIDN:AAA88105.1; PID:g331212
 R:Davidson, A.J.
 Virology 186, 9-14, 1992
 A:Title: Channel catfish virus: a new type of herpesvirus.
 A:Reference number: A39447; MUID:92087490; PMID:1727613
 A:Contents: annotation
 A:Note: neither protein nor nucleic acid sequence is given
 C:Genetics:
 A:Gene: 2
 C:Superfamily: ictalurid herpesvirus 1 hypothetical protein 2

Query Match 57.6%; Score 34; DB 2; Length 407;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 RCRGDDS 9
|||
270 RCLGDDS 276

Db

RESULT 47
F96743
probable C2H2-type zinc finger protein F28P5.6 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: F96743
R:Theologas, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: F96743
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-413 <STO>
A:Cross-references: UNIPROT:Q9C7G8; GB:AE005173; NID:gl1054410; PIDN:AAG27797.1; GSPDB:C
C:Genetics:
A:Gene: F28P5.6
A:Map position: 1

Query Match 57.6%; Score 34; DB 2; Length 413;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 CRGDDSK 10
|:|:|
191 CKGDDDK 197

Db

RESULT 48
H71400
hypothetical protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
A:Variety: columbia
C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 09-Jul-2004
C:Accession: H71400
R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzensegger, T.; Pohl, T.M.; Terry, N.; Giel
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech
erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ans
C.; Chludzisz, N.
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal
A:Reference number: A71400; MUID:98121113; PMID:9461215
A:Accession: H71400
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-482 <BEV>
A:Cross-references: UNIPROT:O23260; GB:297335; NID:g2244747; PID:g2244755
C:Genetics:
A:Map position: 4COP9-4C3845
C:Superfamily: Arabidopsis thaliana hypothetical protein At2g32740

Query Match 57.6%; Score 34; DB 2; Length 482;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYRCRGDDS 9
|||
163 NYRCLTNS 171

Db

RESULT 49
B41627
furin (EC 3.4.21.75) 18 precursor - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 26-May-2000
C:Accession: B41627
R:Korner, J.; Chun, J.; O'Bryan, L.; Axel, R.
Proc. Natl. Acad. Sci. U.S.A. 88, 11393-11397, 1991
A:Title: Prohormone processing in Xenopus oocytes: characterization of cleavage signals
A:Reference number: A41627; MUID:92107956; PMID:1722329
A:Accession: B41627
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-596 <KOR>
A:Cross-references: GB:M80471
C:Superfamily: Kexin; subtilisin homology
C:Keywords: hydrolase; serine proteinase
F:142-380/Domain: subtilisin homology <SBT>
F:151,192,366/Active site: Asp, His, Ser #status predicted
Query Match 57.6%; Score 34; DB 2; Length 596;
Best Local Similarity 71.4%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGD 7
:|:|:|
492 SYNCRGD 498

Db

RESULT 50
A41627
furin (EC 3.4.21.75) 1 precursor - African clawed frog
N:Alternate names: furin homolog 14
C:Species: Xenopus laevis (African clawed frog)
C:Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 09-Jul-2004
C:Accession: A41627
R:Korner, J.; Chun, J.; O'Bryan, L.; Axel, R.
Proc. Natl. Acad. Sci. U.S.A. 88, 11393-11397, 1991
A:Title: Prohormone processing in Xenopus oocytes: characterization of cleavage signals
A:Reference number: A41627; MUID:92107956; PMID:1722329
A:Accession: A41627
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-783 <KOR>
A:Cross-references: UNIPROT:P29119; GB:M80471; NID:g214158; PIDN:AAA49717.1; PID:g214159
C:Superfamily: Kexin; subtilisin homology
C:Keywords: hydrolase; serine proteinase; transmembrane protein
F:142-380/Domain: subtilisin homology <SBT>
F:151,192,366/Active site: Asp, His, Ser #status predicted
Query Match 57.6%; Score 34; DB 2; Length 783;
Best Local Similarity 71.4%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGD 7
:|:|:|
492 SYNCRGD 498

Db

Search completed: September 7, 2005, 20:03:56
Job time : 19.6744 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2005, 19:43:14 ; Search time 35.1163 Seconds
(without alignments)
145.824 Million cell updates/sec

Title: US-10-812-238B-20

Perfect score: 59

Sequence: 1 NYRCRGDDSK 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database : Uniprot_01.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59	100.0	311	1 LPP3_HUMAN	O14495 h lipid pho
2	55	93.2	312	1 LPP3_MOUSE	Q99JY8 mus musculus
3	55	93.2	312	1 LPP3_RAT	P97544 rattus norv
4	55	93.2	312	2 Q6IMX4	Q6IMX4 rattus norv
5	45	76.3	776	2 Q73RY0	Q73RY0 mycobacteri
6	42	71.2	73	2 Q7OP19	Q7OP19 giardia lam
7	42	71.2	265	2 Q86KQ7	Q86KQ7 dictyosteli
8	41	69.5	428	2 Q8TBF1	Q8TBF1 homo sapien
9	41	69.5	489	2 Q8N9B8	Q8N9B8 homo sapien
10	41	69.5	1334	2 Q97322	Q97322 plasmodium
11	40	67.8	150	2 Q9KCP2	Q9KCP2 bacillus ha
12	40	67.8	399	1 PNCB_ECO57	Q8XDE8 escherichia
13	40	67.8	399	1 PNCB_ECOLI	P18133 escherichia
14	40	67.8	399	1 PNCB_SALTI	Q8Z7Y9 salmonella
15	40	67.8	399	1 PNCB_SALTY	P22253 salmonella
16	40	67.8	400	2 Q7UD27	Q7UD27 shigella fl
17	40	67.8	416	2 Q83LN3	Q83LN3 shigella fl
18	40	67.8	416	2 Q8FJ98	Q8FJ98 escherichia
19	40	67.8	705	2 Q5Q468	Q5Q468 mycobacteri
20	40	67.8	705	2 P96283	P96283 mycobacteri
21	40	67.8	705	2 Q77XK7	Q77XK7 mycobacteri
22	40	67.8	845	2 Q8WQJ3	Q8WQJ3 plasmodium
23	40	67.8	1304	2 Q8WT63	Q8WT63 plasmodium
24	39	66.1	37	2 Q8POM2	Q8POM2 streptococ
25	39	66.1	46	2 Q848W4	Q848W4 bacillus me
26	39	66.1	154	2 Q8BUS2	Q8BUS2 mus musculus
27	39	66.1	179	2 Q20683	Q20683 caenorhabdi
28	39	66.1	588	2 Q70681	Q70681 sugarcane s
29	39	66.1	1420	2 Q9YBN1	Q9YBN1 sugarcane s
30	39	66.1	1615	2 Q7RQM4	Q7RQM4 plasmodium
31	38	64.4	121	2 Q8V163	Q8V163 reovirus sp

32	38	64.4	121	2	Q8V164	Q8V164 reovirus sp
33	38	64.4	121	2	Q8V165	Q8V165 reovirus sp
34	38	64.4	121	2	Q8V166	Q8V166 reovirus sp
35	38	64.4	121	2	Q8V167	Q8V167 reovirus sp
36	38	64.4	121	2	Q8V168	Q8V168 reovirus sp
37	38	64.4	121	2	Q8V169	Q8V169 reovirus sp
38	38	64.4	121	2	Q8V170	Q8V170 reovirus sp
39	38	64.4	121	2	Q8V171	Q8V171 reovirus sp
40	38	64.4	121	2	Q8V172	Q8V172 reovirus sp
41	38	64.4	121	2	Q8V173	Q8V173 reovirus sp
42	38	64.4	121	2	Q8V174	Q8V174 reovirus sp
43	38	64.4	121	2	Q8V175	Q8V175 reovirus sp
44	38	64.4	121	2	Q8V176	Q8V176 reovirus sp
45	38	64.4	121	2	Q8V177	Q8V177 reovirus sp
46	38	64.4	121	2	Q8V178	Q8V178 reovirus sp
47	38	64.4	121	2	Q8V179	Q8V179 reovirus sp
48	38	64.4	121	2	Q8V180	Q8V180 reovirus sp
49	38	64.4	121	2	Q8V181	Q8V181 reovirus sp
50	38	64.4	121	2	Q8V182	Q8V182 reovirus sp
51	38	64.4	121	2	Q8V183	Q8V183 reovirus sp
52	38	64.4	121	2	Q8V184	Q8V184 reovirus sp
53	38	64.4	121	2	Q8V185	Q8V185 reovirus sp
54	38	64.4	121	2	Q8V186	Q8V186 reovirus sp
55	38	64.4	121	2	Q8V187	Q8V187 reovirus sp
56	38	64.4	121	2	Q8V188	Q8V188 reovirus sp
57	38	64.4	121	2	Q8V189	Q8V189 reovirus sp
58	38	64.4	121	2	Q8V190	Q8V190 reovirus sp
59	38	64.4	121	2	Q8V191	Q8V191 reovirus sp
60	38	64.4	121	2	Q8V192	Q8V192 reovirus sp
61	38	64.4	121	2	Q8V193	Q8V193 reovirus sp
62	38	64.4	121	2	Q8V194	Q8V194 reovirus sp
63	38	64.4	121	2	Q8V195	Q8V195 reovirus sp
64	38	64.4	121	2	Q8V196	Q8V196 reovirus sp
65	38	64.4	121	2	Q8V197	Q8V197 reovirus sp
66	38	64.4	121	2	Q8V198	Q8V198 reovirus sp
67	38	64.4	121	2	Q8V199	Q8V199 reovirus sp
68	38	64.4	121	2	Q8V1A0	Q8V1A0 reovirus sp
69	38	64.4	121	2	Q8V1A1	Q8V1A1 reovirus sp
70	38	64.4	121	2	Q8V1A2	Q8V1A2 reovirus sp
71	38	64.4	121	2	Q8V1A3	Q8V1A3 reovirus sp
72	38	64.4	121	2	Q8RJU0	Q8RJU0 reovirus sp
73	38	64.4	130	2	Q9VHG0	Q9VHG0 ginglymosto
74	38	64.4	191	2	Q9AY18	Q9AY18 funaria apo
75	38	64.4	193	2	Q9AY16	Q9AY16 entosthodon
76	38	64.4	298	2	Q7Q232	Q7Q232 anopheles g
77	38	64.4	383	2	Q9GTD8	Q9GTD8 caenorhabdi
78	38	64.4	387	2	Q9GTD7	Q9GTD7 caenorhabdi
79	38	64.4	388	2	Q9GTD6	Q9GTD6 caenorhabdi
80	38	64.4	389	2	Q9XTG0	Q9XTG0 caenorhabdi
81	38	64.4	435	2	Q8UTB2	Q8JTB2 golden ide
82	38	64.4	505	2	Q96I80	Q96I80 homo sapien
83	38	64.4	661	2	Q8JTB6	Q8JTB6 striped bas
84	38	64.4	721	2	Q9PA43	Q9PA43 xylella fas
85	38	64.4	875	1	NETR_HUMAN	P56730 homo sapien
86	38	64.4	1240	2	Q8VA42	Q8VA42 chum salmon
87	38	64.4	1267	1	VL3_REOVD	P17378 reovirus ty
88	38	64.4	1267	1	VL3_REOVJ	P17377 reovirus ty
89	38	64.4	1267	1	VL3_REOVL	P17376 reovirus ty
90	38	64.4	1267	2	Q8V5E5	Q8V5E5 ndelle viru
91	38	64.4	1267	2	Q85665	Q85665 reovirus ty
92	38	64.4	1274	2	Q8JU61	Q8JU61 golden shin
93	38	64.4	1274	2	Q9E3V9	Q9E3V9 grass carp
94	37	62.7	59	2	Q6ETU2	Q6ETU2 oryza sativ
95	37	62.7	187	2	Q8HU07	Q8HU07 rhodobryum
96	37	62.7	269	2	Q7R8M3	Q7R8M3 plasmodium
97	37	62.7	307	2	Q86F34	Q86F34 xenopus lae
98	37	62.7	330	2	Q82MF0	Q82MF0 streptomyce
99	37	62.7	430	1	SY5_SYNY3	P73201 synechocyst
100	37	62.7	504	1	Y787_METJA	Q58197 methanococc
101	37	62.7	1268	2	Q63623	Q63623 rattus norv
102	37	62.7	1268	2	Q6DID3	Q6DID3 mus musculu
103	37	62.7	1271	1	RMG_HUMAN	Q9upn6 homo sapien
104	37	62.7	1271	2	Q6NSK3	Q6NSK3 homo sapien

105	37	62.7	1362	2	Q80TJ3	Q80tj3 mus musculus	178	35	59.3	476	2	Q6B8Q6	Q6b8q6 gracilaria
106	37	62.7	4569	2	Q7PS35	Q7ps35 anopheles g	179	35	59.3	503	2	Q6BK51	Q6bk51 debaryomyce
107	36	61.0	63	2	Q9BS83	Q9bs83 homo sapien	180	35	59.3	513	2	Q9A731	Q9a731 caulobacter
108	36	61.0	188	2	Q63XE8	Q63xe8 burkholderi	181	35	59.3	562	2	Q6ZMG3	Q6zmg3 homo sapien
109	36	61.0	226	2	Q69XV4	Q69xv4 oryza sativ	182	35	59.3	575	2	Q8WTU2	Q8wtu2 homo sapien
110	36	61.0	248	2	Q621I5	Q621i5 burkholderi	183	35	59.3	586	2	Q62YB1	Q62yb1 bacillus li
111	36	61.0	248	2	Q63RBD3	Q63rbd3 burkholderi	184	35	59.3	587	2	Q65MW8	Q65mw8 bacillus li
112	36	61.0	266	2	Q8EJ20	Q8ej20 shewanella	185	35	59.3	656	2	Q9M092	Q9m092 arabidopsis
113	36	61.0	306	2	Q7RIN6	Q7rin6 plasmodium	186	35	59.3	662	1	I12R_HUMAN	P42701 homo sapien
114	36	61.0	340	2	Q8V286	Q8v286 helicoverpa	187	35	59.3	757	2	Q6LTA6	Q6lta6 photobacter
115	36	61.0	340	2	Q77LV8	Q77lv8 helicoverpa	188	35	59.3	782	2	Q8PP65	Q8pp65 xanthomonas
116	36	61.0	340	2	Q9E219	Q9e219 helicoverpa	189	35	59.3	849	2	Q815V7	Q815v7 plasmodium
117	36	61.0	340	2	Q91FJ4	Q91fj4 helicoverpa	190	35	59.3	915	2	Q7UV66	Q7uv66 rhodospirell
118	36	61.0	347	2	Q9XAN5	Q9xan5 streptomyce	191	35	59.3	1019	2	Q6CV73	Q6cv73 kluyveromyc
119	36	61.0	401	2	Q6D454	Q6d454 erwinia car	192	35	59.3	1323	2	O13686	O13686 schizosacch
120	36	61.0	404	2	Q7N621	Q7n621 photorhabdu	193	35	59.3	1402	2	Q7RBQ6	Q7rbq6 plasmodium
121	36	61.0	461	2	Q29377	Q29377 archaeoglob	194	35	59.3	1499	2	Q9Y839	Q9y839 mycosphaere
122	36	61.0	461	2	Q69XV5	Q69xv5 oryza sativ	195	35	59.3	1559	2	Q8IDC3	Q8idc3 plasmodium
123	36	61.0	466	2	Q9S5X4	Q9s5x4 thermotoga	196	35	59.3	2736	2	Q7RTC3	Q7rtc3 plasmodium
124	36	61.0	470	2	Q9WYR5	Q9wyr5 thermotoga	197	34.5	58.5	325	2	Q9PP25	Q9pp25 campylobact
125	36	61.0	485	2	Q8CFK1	Q8cfk1 mus musculus	198	34.5	58.5	555	2	Q8H5J9	Q8h5j9 oryza sativ
126	36	61.0	488	2	Q75DX4	Q75dx4 ashbya goss	199	34	57.6	72	2	Q26942	Q26942 triboolium c
127	36	61.0	498	1	NMT2_HUMAN	Q60551 homo sapien	200	34	57.6	87	2	Q9P2Z4	Q9p2z4 homo sapien
128	36	61.0	503	2	Q6FZW1	Q6fzw1 bartonella	201	34	57.6	108	2	Q6ESN2	Q6esn2 oryza sativ
129	36	61.0	503	2	Q6G3B9	Q6g3b9 bartonella	202	34	57.6	112	2	O50739	O50739 borrelia bu
130	36	61.0	507	2	Q8YGM3	Q8ygm3 brucella me	203	34	57.6	122	2	Q40738	Q40738 oryza sativ
131	36	61.0	507	2	Q8G100	Q8g100 brucella su	204	34	57.6	137	2	Q8NAF9	Q8naf9 homo sapien
132	36	61.0	523	2	Q6K632	Q6k632 oryza sativ	205	34	57.6	146	2	Q9CTI3	Q9cti3 mus musculus
133	36	61.0	529	1	NMT2_MOUSE	Q70011 mus musculus	206	34	57.6	164	2	Q7Q0D5	Q7q0d5 anopheles g
134	36	61.0	529	1	Q700Q7	Q700q7 rattus norv	207	34	57.6	164	2	Q8HRE9	Q8hre9 bryum gemmi
135	36	61.0	609	2	Q8IER1	Q8ier1 plasmodium	208	34	57.6	167	2	Q95L84	Q95l84 cryptolagus
136	36	61.0	808	2	Q7O4C6	Q7o4c6 anopheles g	209	34	57.6	172	1	IF3_CAMJEE	Q9p1s2 campylobact
137	36	61.0	897	2	Q93Y95	Q93y95 zea mays (m	210	34	57.6	180	2	O8YJE2	O8yje2 brucella me
138	36	61.0	1022	2	Q8XN34	Q8xn34 clostridium	211	34	57.6	180	2	Q8PYG1	Q8pyg1 brucella su
139	36	61.0	1168	2	Q7RH99	Q7rh99 plasmodium	212	34	57.6	186	2	Q8HRF4	Q8hrf4 ananobryum
140	36	61.0	1363	2	Q81AM5	Q81am5 plasmodium	213	34	57.6	187	2	Q8MG46	Q8mg46 rosulabryum
141	36	61.0	1450	2	Q815X9	Q815x9 plasmodium	214	34	57.6	188	2	Q63005	Q63005 bryum alpin
142	36	61.0	2209	2	Q9U0G6	Q9u0g6 plasmodium	215	34	57.6	188	2	Q9AY14	Q9ay14 physcomitri
143	35.5	60.2	459	2	Q7QJN6	Q7qjn6 anopheles g	216	34	57.6	188	2	O91RL9	O91rl9 arabidopsis
144	35.5	60.2	468	2	Q72WL4	Q72wl4 desulfovibr	217	34	57.6	189	2	O63006	O63006 bryum caesp
145	35	59.3	55	2	Q6SA81	Q6sa81 salmonella	218	34	57.6	189	2	O63011	O63011 rhodobryum
146	35	59.3	128	2	Q74FF9	Q74ff9 geobacter s	219	34	57.6	189	2	O6JE74	O6je74 splachnum w
147	35	59.3	137	2	Q88JY2	Q88jy2 pseudomonas	220	34	57.6	190	2	O62998	O62998 funaria hyg
148	35	59.3	138	2	Q97V09	Q97v09 sulfobolus	221	34	57.6	190	2	O63009	O63009 bryum steno
149	35	59.3	140	2	Q6JLB3	Q6jlb3 francisella	222	34	57.6	190	2	Q8HU00	Q8hu00 bryum palie
150	35	59.3	141	2	Q7MUJ2	Q7muj2 porphyromon	223	34	57.6	190	2	Q8MCT1	Q8mct1 haplodontiu
151	35	59.3	172	2	Q897T4	Q897t4 clostridium	224	34	57.6	190	2	Q9AY13	Q9ay13 aphanorrihag
152	35	59.3	179	2	Q67BF6	Q67bf6 haemophilus	225	34	57.6	191	2	Q9AY15	Q9ay15 physcomitri
153	35	59.3	179	2	Q7VN40	Q7vn40 haemophilus	226	34	57.6	194	2	O63008	O63008 ananobryum
154	35	59.3	187	1	AMPD_SALTY	P30013 salmonella	227	34	57.6	194	2	Q8MCU5	Q8mcu5 bryum russa
155	35	59.3	187	2	Q829F5	Q829f5 salmonella	228	34	57.6	194	2	O8MCU1	O8mcu1 bryum capil
156	35	59.3	190	2	Q96BF5	Q96bf5 homo sapien	229	34	57.6	194	2	O8HRE7	O8hre7 bryum ortho
157	35	59.3	195	2	Q9VRN5	Q9vrn5 drosophila	230	34	57.6	194	2	Q9BCF4	Q9bcf4 physcomitre
158	35	59.3	205	1	BASI_BOVIN	Q865r3 bos taurus	231	34	57.6	195	2	Q8MCU2	Q8mcu2 bryum bicol
159	35	59.3	221	2	Q9ALZ8	Q9alz8 fremyella d	232	34	57.6	195	2	Q8MCU6	Q8mcu6 brachymeni
160	35	59.3	239	2	O44730	Q44730 caenorhabdi	233	34	57.6	196	2	Q8HTZ5	Q8htz5 brachymeni
161	35	59.3	255	2	O86KQ6	Q86kq6 dictyosteli	234	34	57.6	196	2	Q8HU06	Q8hu06 rhodobryum
162	35	59.3	256	2	O86KQ8	Q86kq8 dictyosteli	235	34	57.6	196	2	O8HU13	O8hu13 bryum bland
163	35	59.3	259	2	O6LGM9	Q6lgm9 photobacter	236	34	57.6	196	2	Q8MCT4	Q8mct4 bryum pseud
164	35	59.3	283	2	Q9W057	Q9w057 drosophila	237	34	57.6	196	2	Q8MCU8	Q8mcu8 acidodontiu
165	35	59.3	293	2	Q8RYM8	Q8rym8 oryza sativ	238	34	57.6	197	2	Q8HTZ7	Q8htz7 bryum gemmi
166	35	59.3	304	1	HEM3_XANAC	Q8ppr3 xanthomonas	239	34	57.6	197	2	Q8MTZ8	Q8mtz8 plagibobryum
167	35	59.3	304	1	HEM3_XANCP	Q8p536 xanthomonas	240	34	57.6	197	2	Q8MCT6	Q8mct6 bryum pachy
168	35	59.3	321	2	O9GRJ9	Q9grj9 leishmania	241	34	57.6	197	2	Q8MCT8	Q8mct8 bryum liesae
169	35	59.3	321	2	O8RYK1	Q8ryk1 oryza sativ	242	34	57.6	197	2	Q8MCU0	Q8mcu0 bryum cyath
170	35	59.3	329	1	PIN2_CABEL	Q19157 caenorhabdi	243	34	57.6	197	2	Q8MCU3	Q8mcu3 bryum archa
171	35	59.3	362	2	Q61L91	Q61l91 drosophila	244	34	57.6	197	2	Q8MCU4	Q8mcu4 bryum arach
172	35	59.3	385	1	BASI_HUMAN	P35613 h basigin p	245	34	57.6	197	2	Q8MCU9	Q8mcu9 acidodontiu
173	35	59.3	401	2	Q82N33	Q82n33 streptomyce	246	34	57.6	197	2	Q8MCV0	Q8mcv0 acidodontiu
174	35	59.3	422	2	Q9K426	Q9k426 streptomyce	247	34	57.6	197	2	O8HRF0	O8hrf0 bryum clava
175	35	59.3	439	1	LIA3_HUMAN	Q8n6c8 homo sapien	248	34	57.6	197	2	Q8HRF2	Q8hrf2 brachymeni
176	35	59.3	453	2	P920I3	P920i3 caenorhabdi	249	34	57.6	197	2	Q8HRV8	Q8hrv8 bryum radic
177	35	59.3	467	1	VE2_HPV24	P50770 human papil	250	34	57.6	197	2	Q8HRV9	Q8hrv9 bryum rudr

251	34	57.6	198	2	Q8HU04	Q8hu04 bryum purpu	324	34	57.6	447	2	Q9FDM8	Q9fdm8 streptococc
252	34	57.6	198	2	Q8HU08	Q8hu08 brachymeniu	325	34	57.6	447	2	Q9RPL5	Q9rpl5 streptococc
253	34	57.6	198	2	Q8HU09	Q8hu09 brachymeniu	326	34	57.6	461	2	Q6COC6	Q6coc6 yarrowia li
254	34	57.6	198	2	Q8MCS7	Q8mcs7 rhodobryum	327	34	57.6	470	2	Q755U1	Q755u1 ashbya goss
255	34	57.6	198	2	Q8MCS9	Q8mcs9 bryum cauca	328	34	57.6	482	2	Q232U0	Q232u0 arabidopsis
256	34	57.6	198	2	Q8MCU5	Q8mcu5 bryum algov	329	34	57.6	490	2	Q82137	Q82137 streptomyc
257	34	57.6	198	2	Q8MCU7	Q8mcu7 anomobryum	330	34	57.6	491	2	Q7PPD6	Q7ppd6 anopheles g
258	34	57.6	198	2	Q8HRE8	Q8hre8 bryum coron	331	34	57.6	497	2	Q7SKY5	Q7sky5 oryza sativ
259	34	57.6	199	2	Q63007	Q63007 bryum donia	332	34	57.6	521	2	Q64R16	Q64r16 arabidopsis
260	34	57.6	199	2	Q63010	Q63010 brachymeniu	333	34	57.6	522	2	Q64XQ95	Q64x95 enterobacte
261	34	57.6	199	2	Q8HTZ6	Q8htz6 bryum muehl	334	34	57.6	530	2	Q6BQJ5	Q6bj5 debaryomyce
262	34	57.6	199	2	Q8HU01	Q8hu01 bryum funck	335	34	57.6	593	2	Q86J09	Q86j09 dictyoasteli
263	34	57.6	199	2	Q8HU02	Q8hu02 bryum cellu	336	34	57.6	595	2	Q91751	Q91751 xenopus lae
264	34	57.6	199	2	Q8HU10	Q8hu10 brachymeniu	337	34	57.6	622	2	Q653D2	Q653d2 oryza sativ
265	34	57.6	199	2	Q8HU11	Q8hu11 brachymeniu	338	34	57.6	639	2	Q8D1D1	Q8d1d1 synecococc
266	34	57.6	199	2	Q8HU12	Q8hu12 brachymeniu	339	34	57.6	644	2	Q7MYZ2	Q7myz2 photorhabdu
267	34	57.6	199	2	Q8HU14	Q8hu14 bryum argen	340	34	57.6	770	2	Q653D1	Q653d1 oryza sativ
268	34	57.6	199	2	Q8HRF1	Q8hrf1 brachymeniu	341	34	57.6	783	1	FUR1_XENLA	P29119 xenopus lae
269	34	57.6	199	2	Q8HRF3	Q8hrf3 anomobryum	342	34	57.6	783	2	Q6DE77	Q6de77 xenopus lae
270	34	57.6	199	2	Q9GGF1	Q9gff1 bryum billa	343	34	57.6	801	1	MNHA_STAAM	P60674 staphylococ
271	34	57.6	201	1	RR4_FUNHY	Q9smg1 funaria hyg	344	34	57.6	801	1	MNHA_STAAN	P60675 staphylococ
272	34	57.6	201	1	RR4_PHYPA	Q6yxp3 physcomitre	345	34	57.6	801	1	MNHA_STAAN	Q92ng6 staphylococ
273	34	57.6	202	1	RR4_BRICA	P59135 bryum capil	346	34	57.6	801	1	MNHA_STAAN	Q8nxf6 staphylococ
274	34	57.6	202	2	Q9LRL0	Q9lrl0 arabidopsis	347	34	57.6	801	2	Q6GAX4	Q6gax4 staphylococ
275	34	57.6	203	2	Q6QPD0	Q6qp0 simian aden	348	34	57.6	801	2	Q6GID6	Q6gid6 staphylococ
276	34	57.6	204	2	Q6QP94	Q6qp94 simian aden	349	34	57.6	808	2	Q8N270	Q8n270 homo sapien
277	34	57.6	204	2	Q6QPG6	Q6qp96 simian aden	350	34	57.6	850	1	STB2_YEAST	P46679 saccharomyc
278	34	57.6	250	2	Q6AY82	Q6ay82 rattus norv	351	34	57.6	884	2	Q94J11	Q94j11 oryza sativ
279	34	57.6	254	2	Q7W386	Q7w386 bordetella	352	34	57.6	908	2	Q9FNN3	Q9fn3 arabidopsis
280	34	57.6	254	2	Q7WEK5	Q7wek5 bordetella	353	34	57.6	966	2	Q9LMS1	Q9lms1 arabidopsis
281	34	57.6	255	1	Y166_METJA	Q57630 methanococc	354	34	57.6	970	2	Q7X8K7	Q7x8k7 brassica ju
282	34	57.6	262	2	Q9LRR9	Q9lrr9 arabidopsis	355	34	57.6	970	2	Q9NC89	Q9nc89 strongyloce
283	34	57.6	269	2	Q759L0	Q759l0 ashbya goss	356	34	57.6	988	2	Q9LM32	Q9lm32 arabidopsis
284	34	57.6	270	2	Q9CAB9	Q9cab9 arabidopsis	357	34	57.6	992	2	Q8LPH3	Q8lph3 arabidopsis
285	34	57.6	273	2	Q7XU16	Q7xul6 oryza sativ	358	34	57.6	1030	2	Q8LC25	Q8lc25 arabidopsis
286	34	57.6	273	2	Q9SVH9	Q9svh9 arabidopsis	359	34	57.6	1030	2	Q9C9U5	Q9c9u5 arabidopsis
287	34	57.6	278	2	Q6DHF1	Q6dhf1 brachydanio	360	34	57.6	1036	1	PLD1_CRIGR	O08684 cricetus
288	34	57.6	279	2	Q80V15	Q80v15 mus muscucu	361	34	57.6	1074	1	PLD1_HUMAN	Q13393 homo sapien
289	34	57.6	280	2	Q95698	Q95698 homo sapien	362	34	57.6	1075	2	Q9NC90	Q9nc90 strongyloce
290	34	57.6	280	2	Q9H160	Q9h160 homo sapien	363	34	57.6	1129	2	Q7UKD4	Q7ukd4 rhodopirell
291	34	57.6	281	2	Q856E2	Q856e2 mycobacteri	364	34	57.6	1164	2	Q6SLC8	Q6slc8 cochliobolu
292	34	57.6	281	2	Q8BGU8	Q8bgu8 m mus muscu	365	34	57.6	1233	1	RGY2_SULTO	Q975p6 sulfolobus
293	34	57.6	281	2	Q9ESK4	Q9esk4 mus muscucu	366	34	57.6	1245	2	Q9Y7V5	Q9y7v5 trichoderma
294	34	57.6	288	2	Q9ZU94	Q9zu94 arabidopsis	367	34	57.6	1247	2	Q8I7Z6	Q8i7z6 toxoplasma
295	34	57.6	288	2	Q9DW23	Q9dw23 rat cytomeg	368	34	57.6	1335	2	Q8VI17	Q8vi17 mus muscucu
296	34	57.6	289	2	Q8LDG0	Q8ldg0 arabidopsis	369	34	57.6	1336	2	Q9ES55	Q9es55 mus muscucu
297	34	57.6	296	2	Q92XS2	Q92xs2 rhizobium m	370	34	57.6	2020	2	Q7RQO6	Q7rqg6 plasmodium
298	34	57.6	297	2	Q6H663	Q6h663 oryza sativ	371	34	57.6	2443	2	Q96J17	Q96j17 homo sapien
299	34	57.6	299	2	Q8LA52	Q8las2 arabidopsis	372	34	57.6	2798	2	Q7RPL8	Q7rpl8 plasmodium
300	34	57.6	299	2	Q8GUJ2	Q8guj2 arabidopsis	373	34	57.6	4699	2	Q9V383	Q9v383 drosophila
301	34	57.6	307	2	Q7XIW8	Q7xiw8 oryza sativ	374	33	55.9	82	2	Q8MVZ2	Q8mvz2 ornithodor
302	34	57.6	307	2	Q6AX87	Q6ax87 xenopus lae	375	33	55.9	82	2	Q8MVZ3	Q8mvz3 ornithodor
303	34	57.6	308	2	Q9CQ18	Q9cq18 m mus muscu	376	33	55.9	93	2	Q89U57	Q89us7 bradyrhizob
304	34	57.6	308	2	Q8R220	Q8r220 mus muscucu	377	33	55.9	110	2	Q9DWM8	Q9dwm8 hepatitis c
305	34	57.6	323	2	Q7R2D9	Q7r2d9 giardia lam	378	33	55.9	113	2	Q8H6L0	Q8h6l0 brassica ra
306	34	57.6	324	2	Q940P2	Q940p2 arabidopsis	379	33	55.9	118	2	Q6T399	Q6t399 tomato leaf
307	34	57.6	333	2	Q6A898	Q6a898 propionibac	380	33	55.9	127	2	Q6T886	Q6t886 tomato leaf
308	34	57.6	348	2	Q9SUX6	Q9sux6 arabidopsis	381	33	55.9	127	2	Q8AW35	Q8aw35 brachydanio
309	34	57.6	350	1	SNY_MYCGA	Q7nat8 mycoplasma	382	33	55.9	140	2	Q70K60	Q70k60 gordonia we
310	34	57.6	362	1	CN25_HUMAN	Q7z5m8 homo sapien	383	33	55.9	152	2	Q8ZAU7	Q8zau7 versinia pe
311	34	57.6	372	2	Q9PYT1	Q9pyt1 xestia c-ni	384	33	55.9	154	2	Q7N040	Q7n040 photorhabdu
312	34	57.6	375	2	Q9SYD4	Q9syd4 arabidopsis	385	33	55.9	157	2	Q665H4	Q665h4 versinia pe
313	34	57.6	381	2	Q98C19	Q98c19 rhizobium l	386	33	55.9	157	2	Q8D1Q1	Q8d1q1 versinia pe
314	34	57.6	386	2	Q8X0R8	Q8x0r8 neurospora	387	33	55.9	159	2	Q8TTR5	Q8ttr5 methanoscac
315	34	57.6	392	2	Q86AD7	Q86ad7 dictyosteli	388	33	55.9	167	2	Q9XII2	Q9xii2 arabidopsis
316	34	57.6	393	2	Q9ZVZ4	Q9zvz4 arabidopsis	389	33	55.9	173	2	Q8EIH1	Q8eih1 shewanella
317	34	57.6	407	1	VG02_ICHV1	Q00126 ictalurid h	390	33	55.9	182	2	Q7UST8	Q7ust8 rhodopirell
318	34	57.6	412	2	Q84M24	Q84m24 arabidopsis	391	33	55.9	185	1	RRF_MYCLE	Q33046 mycobacteri
319	34	57.6	413	2	Q9C7G8	Q9c7g8 arabidopsis	392	33	55.9	187	2	Q68IK5	Q68ik5 helicobacte
320	34	57.6	421	2	O44635	O44635 caenorhabdi	393	33	55.9	196	2	Q8MCS8	Q8mcs8 rhodobryum
321	34	57.6	429	1	SYS_GLOVI	Q7nmp9 gloeobacter	394	33	55.9	200	2	O02498	O02498 leishmania
322	34	57.6	437	2	Q6FQ58	Q6fq58 candida gla	395	33	55.9	204	2	Q7M5K8	Q7m5k8 simian aden
323	34	57.6	447	2	Q9F0G1	Q9f0g1 streptococc	396	33	55.9	217	2	Q8L5E8	Q8l5e8 hordeum vul

397	33	55.9	226	1	YV39_ANASP	Q8yee9 anabaena sp	470	33	55.9	559	2	Q9E7C0	Q9e7c0 cucumber ve
398	33	55.9	226	2	Q73222	Q73222 mycobacteri	471	33	55.9	566	2	Q8S313	Q8s313 phytophthor
399	33	55.9	228	2	Q33530	Q33530 rhizobium l	472	33	55.9	567	2	Q8SRV0	Q8srvo encephalito
400	33	55.9	238	2	Q74551	Q74551 schizosacch	473	33	55.9	569	2	Q9M0C9	Q9m0c9 arabidopsis
401	33	55.9	244	2	Q93WH5	Q93wh5 oryza sativ	474	33	55.9	571	2	Q8X003	Q8x003 neurospora
402	33	55.9	248	1	YDFG_ECOLI	P39831 escherichia	475	33	55.9	574	2	Q73MU5	Q73mu5 treponema d
403	33	55.9	248	2	Q65S85	Q65s85 manheimia	476	33	55.9	576	2	Q7RVN7	Q7rvn7 neurospora
404	33	55.9	248	2	Q7UCH2	Q7uch2 shigella fl	477	33	55.9	577	1	QY4B_CAEEL	Q17426 caenorhabdi
405	33	55.9	248	2	Q8X505	Q8x505 escherichia	478	33	55.9	579	2	Q67Y17	Q67y17 arabidopsis
406	33	55.9	253	2	Q83RE8	Q83re8 shigella fl	479	33	55.9	581	2	Q8Y5M8	Q8y5m8 mus musculu
407	33	55.9	253	2	Q8FHD2	Q8fhd2 escherichia	480	33	55.9	581	2	Q8S660	Q8s660 barley mild
408	33	55.9	255	2	Q8MI30	Q8mi30 equus caball	481	33	55.9	582	2	Q911I2	Q911i2 cardanom mo
409	33	55.9	257	2	Q91RK3	Q91rk3 arabidopsis	482	33	55.9	584	2	Q517I7	Q517i7 pseudomonas
410	33	55.9	259	2	Q7V949	Q7v949 prochloroco	483	33	55.9	600	1	SVR_CAUCR	Q9a347 caulobacter
411	33	55.9	270	2	Q70H83	Q70h83 fowlpox vir	484	33	55.9	605	2	Q6BK98	Q6bk98 debaryomyce
412	33	55.9	270	2	Q9J5E2	Q9j5e2 fowlpox vir	485	33	55.9	608	2	Q73JW5	Q73jw5 treponema d
413	33	55.9	273	2	Q9JUV3	Q9juv3 neisseria m	486	33	55.9	625	2	Q9D3V8	Q9d3v8 mus musculu
414	33	55.9	287	2	P72287	P72287 rhizobium l	487	33	55.9	640	2	Q8QS61	Q8qs61 pongine her
415	33	55.9	295	2	Q7NGY7	Q7ngy7 gloeobacter	488	33	55.9	644	2	Q8A3E1	Q8a3e1 bacteroides
416	33	55.9	299	2	Q6GL93	Q6gl93 xenopus tro	489	33	55.9	657	2	Q14006	Q14006 homo sapien
417	33	55.9	307	2	Q6ZLA2	Q6zla2 oryza sativ	490	33	55.9	658	2	Q8P3D2	Q8p3d2 xanthomonas
418	33	55.9	309	2	Q6INT6	Q6int6 xenopus lae	491	33	55.9	658	2	Q8PEQ9	Q8peq9 xanthomonas
419	33	55.9	310	2	Q8DIT7	Q8dit7 synechococc	492	33	55.9	658	2	Q87928	Q87928 xylella fas
420	33	55.9	312	2	Q66KV0	Q66kv0 xenopus lae	493	33	55.9	667	2	Q6FWT7	Q6fwt7 candida glia
421	33	55.9	323	2	Q8EM66	Q8em66 oceanobacil	494	33	55.9	675	2	Q55177	Q55177 synechocyst
422	33	55.9	329	2	Q9NHE0	Q9nhe0 meloidogyne	495	33	55.9	678	2	Q9QBG6	Q9qbg6 wheat yello
423	33	55.9	336	1	HUTG_VIBCH	Q9kaq2 vibrio chol	496	33	55.9	680	2	Q9PA40	Q9pa40 xylella fas
424	33	55.9	347	1	COBT_VIBPA	Q87q46 vibrio para	497	33	55.9	707	2	Q8CD82	Q8cd82 mycobacteri
425	33	55.9	352	2	Q9SYB1	Q9syb1 arabidopsis	498	33	55.9	719	2	P87541	P87541 barley mild
426	33	55.9	360	2	Q9LKY3	Q9lky3 glycine max	499	33	55.9	725	2	Q73634	Q73634 xenopus lae
427	33	55.9	360	2	Q9LKY4	Q9lky4 glycine max	500	33	55.9	727	2	Q6RKB2	Q6rkb2 rattus norv
428	33	55.9	370	2	Q8LL17	Q8ll17 arabidopsis							
429	33	55.9	389	1	HIS8_CANMA	P56099 candida mal							
430	33	55.9	399	2	Q90BG5	Q9qbg5 wheat yello							
431	33	55.9	401	2	Q7S9U6	Q7s9u6 neurospora							
432	33	55.9	403	2	Q6BMG0	Q6bm90 debaryomyce							
433	33	55.9	411	2	Q8SXG0	Q8sxg0 drosophila							
434	33	55.9	412	2	Q9VX69	Q9vx69 anopheles g							
435	33	55.9	412	2	Q7Q1R5	Q7qlr5 anopheles g							
436	33	55.9	413	2	Q6REL5	Q6rel5 rhodococcus							
437	33	55.9	419	2	Q6GLS6	Q6gl86 xenopus lae							
438	33	55.9	420	2	Q6LNB3	Q6lnb3 photobacter							
439	33	55.9	421	2	Q8YF04	Q8yf04 bruceella su							
440	33	55.9	421	2	Q8G2T7	Q8g2t7 bruceella su							
441	33	55.9	421	2	Q41898	Q41898 rice necros							
442	33	55.9	432	2	Q98LP4	Q98lp4 rhizobium l							
443	33	55.9	448	2	Q91BP9	Q91bp9 oat mosaic							
444	33	55.9	460	2	Q9N8R2	Q9n8r2 trypanosoma							
445	33	55.9	469	2	Q6Z4U6	Q6z4u6 oryza sativ							
446	33	55.9	481	2	Q7MXE2	Q7mxe2 porphyromon							
447	33	55.9	485	2	Q8P748	Q8p748 xanthomonas							
448	33	55.9	488	2	Q8X0T4	Q8x0t4 mus musculu							
449	33	55.9	488	2	Q91QR6	Q91qr6 wheat yello							
450	33	55.9	489	2	Q9A762	Q9a762 caulobacter							
451	33	55.9	490	2	Q9BM62	Q9bm62 homo sapien							
452	33	55.9	494	2	Q9Z012	Q9z012 wheat spin							
453	33	55.9	498	2	Q89M54	Q89m54 bradyrhizob							
454	33	55.9	505	2	Q98NR7	Q98nr7 rhizobium l							
455	33	55.9	508	2	Q78617	Q78617 neurospora							
456	33	55.9	510	2	Q6M0Z3	Q6m0z3 methanococc							
457	33	55.9	513	2	Q83869	Q83869 narcisus l							
458	33	55.9	514	2	Q80BV2	Q80bv2 alpinia mos							
459	33	55.9	528	2	Q97379	Q97379 strongyloce							
460	33	55.9	528	2	Q6J0M7	Q6jom7 indian card							
461	33	55.9	528	2	Q9QD49	Q9qd49 cardanom mo							
462	33	55.9	540	2	Q76771	Q76771 dictyosteli							
463	33	55.9	543	2	Q17900	Q17900 caenorhabdi							
464	33	55.9	544	2	Q7T810	Q7t810 wheat strei							
465	33	55.9	545	2	Q9U3Z2	Q9u3z2 hydra atten							
466	33	55.9	546	2	Q29111	Q29111 sus scrofa							
467	33	55.9	554	2	Q7TF98	Q7tf98 cucumber ve							
468	33	55.9	555	2	Q9QB93	Q9qb93 wheat yello							
469	33	55.9	557	2	Q91DH6	Q91dh6 chinese yam							

ALIGNMENTS

RESULT 1

LP3_HUMAN	STANDARD;	PRT;	311 AA.
ID_LPP3_HUMAN	Q96GW0; Q99782;		
AC	Q14495; Q96GW0; Q99782;		
DT	29-MAR-2004 (Rel. 43, Created)		
DT	29-MAR-2004 (Rel. 43, Last sequence update)		
DT	05-JUL-2004 (Rel. 44, Last annotation update)		
DE	Lipid phosphate phosphohydrolase 3 (EC 3.1.3.4) (Phosphatidic acid		
DE	phosphatase 2b) (Phosphatidate phosphohydrolase type 2b) (PAP2b) (PAP-		
DE	2b) (PAP2-beta) (Vascular endothelial growth factor and type I		
DE	collagen inducible protein) (VCIP).		
GN	Name=PAP2b; Synonyms=LPP3;		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A., CHARACTERIZATION, AND INDUCTION.		
RX	MEDLINE=97450990; PubMed=9305923; DOI=10.1074/jbc.272.39.24572;		
RA	Kai M., Wada I., Imai S.-I., Sakane F., Kanoh H.;		
RT	"Cloning and characterization of two human isozymes of Mg2+-		
RT	independent phosphatidic acid phosphatase.";		
RL	J. Biol. Chem. 272:24572-24578(1997).		
RN	[2]		
RP	SEQUENCE FROM N.A., AND CHARACTERIZATION.		
RX	MEDLINE=98371049; PubMed=9705349; DOI=10.1074/jbc.273.34.22059;		
RA	Roberts R., Sciorra V.A., Morris A.J.;		
RT	"Human type 2 phosphatidic acid phosphohydrolases. Substrate		
RT	specificity of the type 2a, 2b, and 2c enzymes and cell surface		
RT	activity of the 2a isoform.";		
RL	J. Biol. Chem. 273:22059-22067(1998).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RX	PubMed=12660161; DOI=10.1093/emboj/cdg165;		
RA	Huntsree J.O., Feng S., Thakker G.D., Yang J., Hong J., Wary K.K.;		
RT	"Regulation of cell-cell interactions by phosphatidic acid phosphatase		
RT	2b/VCIP.";		

RL EMBL J. 22:1539-1554 (2003).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Leung D.W., Tompkins C.K.;
 RT "Molecular cloning of and expression of an isoform of human
 RA phosphatidic acid phosphatase cDNA";
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=97264341; PubMed=9110174;
 RX Yu W., Anderson B., Worley K.C., Muzny D.M., Ding Y., Liu W.,
 RA Ricafrente J.Y., Wentland M.A., Lennon G., Gibbs R.A.;
 RT "Large-scale concatenation cDNA sequencing";
 RL Genome Res. 7:353-358 (1997).
 RN [6]
 RP SEQUENCE FROM N.A.
 RA TISSUE=Placenta;
 RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udutin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN SUBUNIT.
 RP PubMed=14727515; DOI=10.1186/1471-2091-5-2;
 RX Burnett C., Makridou P., Hewlett L., Howard K.;
 RA "Lipid phosphate phosphatases dimerize, but this interaction is not
 RT required for in vivo activity";
 RL BMC Biochem. 5:2-2 (2004).
 CC -I- FUNCTION: Catalyzes the conversion of phosphatidic acid (PA) to
 CC diacylglycerol (DG). In addition it hydrolyzes lysophosphatidic
 CC acid (LPA), ceramide-1-phosphate (C-1-P) and sphingosine-1-
 CC phosphate (S-1-P). The relative catalytic efficiency is LPA = PA >
 CC C-1-P > S-1-P. May be involved in cell adhesion and in cell-cell
 CC interactions.
 CC -I- CATALYTIC ACTIVITY: A 3-en-phosphatidate + H(2)O = a 1,2-diacyl-
 CC an-glycerol + phosphate.
 CC -I- ENZYME REGULATION: Inhibited by sphingosine, zinc ions and
 CC propanolol. Not inhibited by N-ethylmaleimide treatment.
 CC -I- SUBUNIT: Homodimer. This complex seems not to be involved in
 CC substrate recognition, it may confer only structural or functional
 CC stability.
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein. Post-Golgi and
 CC plasma membrane localization.
 CC -I- TISSUE SPECIFICITY: Ubiquitously expressed. Highly expressed in
 CC heart and placenta.
 CC -I- INDUCTION: By epidermal growth factor (EGF), vascular endothelial
 CC growth factor (VEGF), basic fibroblast growth factor (bFGF) and
 CC phorbol myristate acetate (PMA).
 CC -I- PTM: N-glycosylated. Contains high-mannose oligosaccharides.
 CC -I- SIMILARITY: Belongs to the PA-phosphatase related phosphoesterase
 CC family.
 CC -I- CAUTION: Ref.5 sequence differs from that shown due to a
 CC frameshift in position 225.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB000889; BAA22594.1; -;
 DR EMBL; AF017786; AAC63433.1; -;
 DR EMBL; AF408083; AA084481.1; -;
 DR EMBL; AF043329; AAD02271.1; -;
 DR EMBL; U79294; AAB50222.1; ALT_FRAME.
 DR EMBL; BC009196; AAH09196.1; -;
 DR Genew; HGNC:9229; PPAP2B.
 DR H-InvDB; HIX0000628; -;
 DR Reactome; O14495; -;
 DR MIM; 607125; -;
 DR GO; GO:0016020; C-membrane; TAS.
 DR GO; GO:0004721; P-phosphoprotein phosphatase activity; TAS.
 DR GO; GO:0008151; P-cell growth and/or maintenance; TAS.
 DR GO; GO:0008354; P-germ-cell migration; TAS.
 DR InterPro; IPR008934; AcPase VanPase.
 DR InterPro; IPR000326; Peptidase_PA_PTP.
 DR Pfam; PF01569; PAP2; 1.
 DR SMART; SM00014; acidPPc; 1.
 KW Glycoprotein; Hydrolase; Transmembrane.
 FT DOMAIN 1 33 Cytoplasmic (Potential).
 FT TRANSMEM 34 54 Potential.
 FT DOMAIN 55 85 Lumenal (Potential).
 FT TRANSMEM 86 106 Potential.
 FT DOMAIN 107 122 Cytoplasmic (Potential).
 FT TRANSMEM 123 143 Potential.
 FT DOMAIN 144 193 Lumenal (Potential).
 FT TRANSMEM 194 214 Potential.
 FT DOMAIN 215 227 Cytoplasmic (Potential).
 FT TRANSMEM 228 248 Potential.
 FT DOMAIN 249 257 Lumenal (Potential).
 FT TRANSMEM 258 278 Potential.
 FT DOMAIN 279 311 Cytoplasmic (Potential).
 FT CARBOHYD 170 170 N-linked (GlcNAc...) (Potential).
 FT CONFLICT 282 282 T -> M (in Ref. 6).
 SQ SEQUENCE 311 AA; 35116 MW; CB3F60189044DA31 CRC64;
 Query Match 100.0%; Score 59; DB 1; Length 311;
 Best Local Similarity 100.0%; Pred. No. 0.0084;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NYRCRGDDSK 10
 Db 178 NYRCRGDDSK 187
 RESULT 2
 LPP3 MOUSE
 ID LPP3 MOUSE STANDARD; PRT; 312 AA.
 AC Q9JF8; Q8BTB7;
 DT 29-MAR-2004 (Rel. 43, Created)
 DT 29-MAR-2004 (Rel. 43, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Lipid phosphate phosphohydrolase 3 (EC 3.1.3.4) (Phosphatidic acid
 DE phosphatase 2b) (Phosphatidate phosphohydrolase type 2b) (PAP2b) (PAP-
 DE 2b) (PAP2-beta).
 GN Name=Ppap2b; Synonyms=lpp3;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CS7BL/6J; TISSUE=Embryo;
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,

RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Glissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagaehima T., Numata K., Okido T., Pavan W.J., Perlee G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirose-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Inotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sakaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S.J., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Murzyn D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting J., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RN FUNCTION, AND DISEASE.
RX PubMed=12925589; DOI=10.1242/dev.00635;
RA Escalante-Alcalde D., Hernandez L., Le Stunff H., Maeda R., Lee H.-S.,
RA Cheng G. Jr., Sciorra V.A., Daar I., Spiegel S., Morris A.J.,
RA Stewart C.L.;
RT "The lipid phosphatase LPP3 regulates extra-embryonic vasculogenesis
RT and axis patterning.";
RL Development 130:4623-4637 (2003).
CC -!- FUNCTION: Catalyzes the conversion of phosphatidic acid (PA) to
CC diacylglycerol (DG). In addition it hydrolyzes lysophosphatidic
CC acid (LPA), ceramide-1-phosphate (C-1-P) and sphingosine-1-
CC phosphate (S-1-P) (By similarity). Essential to the formation of
CC the chorio-allantoic placenta and extra-embryonic vasculature.
CC Also mediates gastrulation and axis formation, probably by
CC regulating the Wnt signaling pathway.
CC -!- CATALYTIC ACTIVITY: A 3-sn-phosphatidate + H(2)O = 1,2-diacyl-
CC sn-glycerol + phosphate.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC reticulum associated (By similarity).
CC -!- DISEASE: Ppap2b deficient embryos fail to form a chorio-allantoic
CC placenta and yolk sac vasculature. A subset of embryos also show a
CC shortening of the anterior-posterior axis and frequent duplication

CC of axial structures. Loss of Ppap2b results in a marked increase
CC in beta-catenin-mediated T-cell factor (TCF) transcription.
CC -!- SIMILARITY: Belongs to the PA-phosphatase related phosphoesterase
CC family.
CC -!- CAUTION: Ref.1 sequence differs from that shown due to a
CC frameshift in position 226.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AK011276; BAC35327.1; ALT_FRAME.
CC EMBL; BC055559; AAH05558.1; --.
CC MGD; MGI:1915166; Ppap2b.
CC GO; GO:0042577; F:lipid phosphatase activity; IMP.
CC GO; GO:0001569; P:blood vessel development; IMP.
CC GO; GO:0010003; P:gastrulation (sensu Mammalia); IMP.
CC GO; GO:0006644; P:phospholipid metabolism; IMP.
CC GO; GO:0030111; P:regulation of Wnt receptor signaling pathway; IDA.
CC InterPro; IPR008934; Acpase_Vanperase.
CC InterPro; IPR000326; Pesterase_PA_PTP.
CC Pfam; PF01569; PAP2; 1.
CC SMART; SM00014; acidPPc; 1.
CC Developmental protein; Glycoprotein; Hydrolase; Transmembrane.
CC DOMAIN 1 33 Cytoplasmic (Potential).
CC TRANSMEM 34 54 Potential.
CC DOMAIN 55 85 Luminal (Potential).
CC TRANSMEM 86 106 Potential.
CC DOMAIN 107 123 Cytoplasmic (Potential).
CC TRANSMEM 124 144 Potential.
CC DOMAIN 145 194 Luminal (Potential).
CC TRANSMEM 195 215 Potential.
CC DOMAIN 216 226 Cytoplasmic (Potential).
CC TRANSMEM 227 247 Potential.
CC DOMAIN 248 258 Luminal (Potential).
CC TRANSMEM 259 279 Potential.
CC DOMAIN 280 312 Cytoplasmic (Potential).
CC CARBOHYD 171 171 N-linked (GlcNAc...) (Potential).
CC SEQUENCE 312 AA; 35216 MW; D782986E04B57D7D CRC64;
SQ
Query Match 93.12%; Score 55; DB 1; Length 312;
Best Local Similarity 90.0%; Pred. No. 0.048;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 NYRCRGDSK 10
|||||:|||||
Db 179 NYRCRGDSK 188
RESULT 3
LPP3 RAT ID LPP3 RAT STANDARD; PRT; 312 AA.
AC P97544;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Lipid phosphate phosphohydrolase 3 (EC 3.1.3.4) (Phosphatidic acid
DE phosphate 2b) (Phosphatidate phosphohydrolase type 2b) (PAP2b) (PAP-
DE 2b) (PAP2-beta) (Differentially expressed in rat intestine 42)
DE (Dri42).
GN Name=Ppap2b; Synonyms=lpp3;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., TOPOLOGY, AND N-GLYCOSYLATION.
RC STRAIN=Wistar; TISSUE=Small intestine;
RX MEDLINE=97094703; PubMed=8939937; DOI=10.1074/jbc.271.47.29928;

Qy 1 NTRCRGDSK 10
|||||:||||
Db 179 NTRCRGDSK 188

RESULT 4
Q61MX4 PRELIMINARY; PRT; 312 AA.

AC Q61MX4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE ER transmembrane protein Dri 42.
GN Name=Dri42;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=101116;
RN [1]
SEQUENCE FROM N.A.
RP
RC TISSUE=Heart;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zebberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uddin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worey K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Greenwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Matra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
SEQUENCE FROM N.A.
RP
RC TISSUE=Heart;
RA Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC072544; AAH72544.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR008934; AcPase VanParase.
DR InterPro; IPR000326; Pesterase_PA_PTP.
DR Pfam; PF01569; PAP2; 1
DR SMART; SM00014; acidPPc; 1.
KW Transmembrane.
SQ SEQUENCE 312 AA; 35233 MW; CDA54495C0E7D37D CRC64;

Query Match 93.2%; Score 55; DB 2; Length 312;
Best Local Similarity 90.0%; Pred. No. 0.048;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTRCRGDSK 10
|||||:||||
Db 179 NTRCRGDSK 188

RESULT 5
Q73RY0 PRELIMINARY; PRT; 776 AA.

AC Q73RY0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.

```

GN OrderedLocusNames=MAP4298c;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterinae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AS017242; AAS06848.1; -.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR003673; CAIB_BAIF.
DR Pfam; PF02515; CoA_transf_3; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 776 AA; 82189 MW; 38A4528FC5B0B4BB CRC64;

Query Match 76.3%; Score 45; DB 2; Length 776;
Best Local Similarity 87.5%; Pred. No. 9.2;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDD 9
Db 615 YRCRGDDA 622
|||||:

RESULT 6
Q7QP19
ID Q7QP19 PRELIMINARY; PRT; 73 AA.
AC Q7QP19;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE GLP 83 1716 1495.
OS Giardia lamblia ATCC 50803.
OC Eukaryota; Diplomonadida; Hexamitidae; Giardiinae; Giardia.
OX NCBI_TaxID=184922;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WB C6;
RA Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D.,
RA Olsen G.J., Sogin M.L.;
RL "Draft sequence of the Giardia lamblia genome.";
RT Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACB01000181; EAA36741.1; -.
SQ SEQUENCE 73 AA; 8125 MW; 4F866C8ABF254DDE CRC64;

Query Match 71.2%; Score 42; DB 2; Length 73;
Best Local Similarity 75.0%; Pred. No. 3;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYRCRGDD 8
Db 27 DYRCRGDD 34
|||||:

RESULT 7
Q86KQ7
ID Q86KQ7 PRELIMINARY; PRT; 265 AA.
AC Q86KQ7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;

RX MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lemmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunngal B., Cox B., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
RL Nature 418:79-85(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Baumgart C.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC115584; AAO50949.1; -.
DR InterPro; IPR008997; RicinB_like.
KW Hypothetical protein.
SQ SEQUENCE 265 AA; 30503 MW; F1E238D4A048F6D2 CRC64;

Query Match 71.2%; Score 42; DB 2; Length 265;
Best Local Similarity 87.5%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 YRCRGDD 9
Db 45 YRCRNDD 52
|||||:

RESULT 8
Q8TBF1
ID Q8TBF1 PRELIMINARY; PRT; 428 AA.
AC Q8TBF1;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE RasGEF domain family, member 1A.
GN Name=RASGEF1A;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=23388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC022548; AAH22548.1; -.
DR GO; GO:0005085; P:guanylnucleotide exchange factor activity; IEA.
DR GO; GO:0007264; P:small GTPase mediated signal transduction; IEA.
DR InterPro; IPR000651; RasGef_N.
DR InterPro; IPR001895; RasGRF_CDC25.
DR InterPro; IPR008937; Ras_GEF.

```

```

DR Pfam: PF00617; RasGEF; 1.
DR SMART: SM00147; RasGEF; 1.
DR PROSITE; PSS0009; RasGEF_CAT; 1.
DR PROSITE; PSS0212; RasGEF_NTER; 1.
SQ SEQUENCE 428 AA; 49394 MW; 2522DBF4A3D2F560 CRC64;

Query Match 69.5%; Score 41; DB 2; Length 428;
Best Local Similarity 70.0%; Pred. No. 28;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDSDK 10
DB 197 NHRCRGDLTK 206
|:|||||:|

RESULT 9
Q8N9B8 PRELIMINARY; PRT; 489 AA.
ID Q8N9B8
AC Q8N9B8
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein FLJ37817.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RX TISSUE=Substantia nigra;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuma T., Sato K., Tanikawa M., Yamazaki M.,
RA Ninomiya K., Ihibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimata T., Watanabe M., Hiraoaka S., Chiba Y., Iehida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Yoshikawa K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Moriya S., Momiya H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takenoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RA "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs";
RL Nat. Genet. 36:40-45(2004).
DR ENBL; AK095136; BAC04491.1; -.
DR DR GO; GO:0005085; P:guanylyl-nucleotide exchange factor activity; IEA.
DR DR GO; GO:0007264; P:small GTPase mediated signal transduction; IEA.
DR DR InterPro; IPR000651; RasGEF_N.
DR DR InterPro; IPR001895; RasGEF_CDC25.
DR DR InterPro; IPR008937; Ras_GEF.
DR DR Pfam; PF00617; RasGEF; 1.
DR DR SMART; SM00147; RasGEF; 1.
DR DR SMART; SM00229; RasGEFN; 1.
DR DR PROSITE; PSS0009; RasGEF_CAT; 1.
DR DR PROSITE; PSS0212; RasGEF_NTER; 1.

SQ SEQUENCE 489 AA; 55532 MW; 85D5986ACD2C7C76 CRC64;

Query Match 69.5%; Score 41; DB 2; Length 489;
Best Local Similarity 70.0%; Pred. No. 32;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDSDK 10
DB 258 NHRCRGDLTK 267
|:|||||:|

RESULT 10
O97322 PRELIMINARY; PRT; 1334 AA.
ID O97322
AC O97322
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein MAL3P7.32
GN Name=MAL3P7.32; Synonyms=PF01010w;
OS Plasmodium falciparum (Isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99376085; PubMed=10448855; DOI=10.1038/22964;
RA Bowman S., Lawson D., Basham D., Brown D., Chillingworth T.,
RA Churcher C.M., Craig A., Davies R.M., Devlin K., Feltwell T.,
RA Gentles S., Gwilliam R., Hamlin N., Harris D., Holroyd S., Hornsby T.,
RA Horrocks P., Jagels K., Jasek B., Kyes S., McLean J., Moule S.,
RA Mungall K., Murphy L., Oliver K., Quail M.A., Rajandream M.-A.,
RA Rutter S., Skelton J., Squares R., Squares S., Sulston J.E.,
RA Whitehead S., Woodward J.R., Newbold C., Barrell B.G.;
RT "The complete nucleotide sequence of chromosome 3 of Plasmodium
RT falciparum";
RL Nature 400:532-538(1999).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22255708; PubMed=12369867; DOI=10.1038/nature01095;
RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
RA Mungall K., Bowman S., Atkin R., Baker S., Barton A., Brooks K.,
RA Buckee C.O., Burrows C., Cherevach I., Chillingworth C., Corton C.,
RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
RA Feltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
RA Humphray S., Jagels K., James K.D., Johnson D., Kerhornou A.,
RA Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,
RA Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,
RA Oliver K., Ormond D., Price C., Quail M.A., Rabinowitch E.,
RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
RA Sulston J.E., Craig A., Newbold C., Barrell B.G.;
RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13.";
RL Nature 419:527-531(2002).
DR ENBL; AL034559; CAB39073.2; -.
RW Hypothetical protein.
SQ SEQUENCE 1334 AA; 160847 MW; E3577E84C7E0C8E5 CRC64;

Query Match 69.5%; Score 41; DB 2; Length 1334;
Best Local Similarity 60.0%; Pred. No. 91;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYRCRGDSDK 10
DB 808 NYNCKDDDNK 817
|:|||||:|

RESULT 11
O9KCP2 PRELIMINARY; PRT; 150 AA.
ID O9KCP2
AC O9KCP2;

```

```

DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Ferric uptake regulation protein.
GN OrderedLocusNames=BH1527;
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125;
RX MEDLINE=20512582; PubMed=11058132; DOI=10.1093/nar/28.21.4317;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331 (2000).
DR EMBL; AF001512; BAB05246.1; -.
DR PIR; G83840; G83840.
DR HSSP; Q03456; 1M2B.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR002481; FUR.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR Pfam; PF01475; FUR; 1.
DR ProDom; PD002003; FUR; 1.
KW Complete Proteome.
SQ SEQUENCE 150 AA; 17447 MW; CE4D3AC2A8B64275 CRC64;

Query Match 67.8%; Score 40; DB 2; Length 150;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDDSK 10
Db 141 HRCQGDSEK 149
:|||||

RESULT 12
PNCB_ECO57 STANDARD; PRT; 399 AA.
AC Q8XDE8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Nicotinate phosphoribosyltransferase (EC 2.4.2.11) (NAPRTase).
GN Escherichia coli O157:H7.
OS Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dinalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / Sakai / RIMD 050952 / EHEC;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shingawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT "Complete genome sequence of Escherichia coli K-12.";
```

```

RT O157:H7 and genomic comparison with a laboratory strain K-12.";
```

DNA Res. 8:11-22(2001).
 CC -!- CATALYTIC ACTIVITY: Nicotinate D-ribose 1-diphosphate + diphosphate =
 CC nicotinate + 5-phospho-alpha-D-ribose 1-diphosphate.
 CC -!- PATHWAY: NAD biosynthesis; nicotinamide to NAMN; second step.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the NAPRTase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AE005283; AAG54416.1; -.
 DR EMBL; AF002553; BAB34437.1; -.
 DR PIR; D85619; D85619.
 DR PIR; F90755; F90755.
 DR HAMAP; MF 00570; -. 1.
 DR InterPro; IPR007229; NAPRTase.
 DR InterPro; IPR006406; Nic.Prtans.
 DR InterPro; IPR008967; P53_like_DNA_bnd.
 DR Pfam; PF04095; NAPRTase; 1.
 DR PIRSF; PIRSF000484; Nicot_phos_ribo; 1.
 DR TIGSFAM; TIGR01514; NAPRTase; 1.
 KW Complete proteome; Glycosyltransferase;
 KW Pyridine nucleotide biosynthesis; Transferase.
 FT INIT MET 0 By similarity.
 SQ SEQUENCE 399 AA; 45838 MW; 47DAC387A0EC6926 CRC64;

Query Match 67.8%; Score 40; DB 1; Length 399;
Best Local Similarity 85.7%; Pred. No. 41;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDD 8
Db 38 FRCRGDD 44
:|||||

RESULT 13
PNCB_ECOLI STANDARD; PRT; 399 AA.
AC P18133;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Nicotinate phosphoribosyltransferase (EC 2.4.2.11) (NAPRTase).
GN Name=pncB; OrderedLocusNames=b0931;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-20.
RX STRAIN=CEC70;
RX MEDLINE=91009224; PubMed=2211655;
RA Wubbolds M.G., Terpstra P., van Beilen J.B., Kingma J.,
RA Meesters H.A.R., Witholt B.;
RT "Variation of cofactor levels in Escherichia coli. Sequence analysis
RT and expression of the pncB gene encoding nicotinic acid
RT phosphoribosyltransferase.";
RL J. Biol. Chem. 265:17665-17672 (1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Colliado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";


```

RL Science 277:1453-1474 (1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayaishi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa K., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
RC
CC -1- CATALYTIC ACTIVITY: Nicotinate D-ribonucleotide + diphosphate =
CC nicotinate + 5-phospho-alpha-D-ribose 1-diphosphate.
CC -1- PATHWAY: NAD biosynthesis; nicotinamide to NAMN; second step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SUBCELLANEQUS: Requires ATP as well as phosphoribosyl
CC pyrophosphate for activity.
CC -1- SIMILARITY: Belongs to the NAPRTase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J05568; AAA24400.1; --
CC EMBL; U00096; AAC74017.1; --
CC EMBL; D90731; BAA35683.1; --
CC PIR; JQ0756; JQ0756.
CC ECHOBASE; EB0735; --
CC RcoGene; EG10742; pncB.
CC HAMAP; MF_00570; --; 1.
CC InterPro; IPR007229; NAPRTase.
CC InterPro; IPR006406; Nic Prtrans.
CC InterPro; IPR008967; P53_like_DNA_bnd.
CC Pfam; PF04095; NAPRTase; 1.
CC PIRSF; PIRSF000484; Nicot_phos_ribo; 1.
CC TIGRFAMs; TIGR01514; NAPRTase; 1.
CC Complete proteome; direct protein sequencing; Glycosyltransferase;
KW Pyridine nucleotide biosynthesis; Transferase.
FT INIT MET 0
FT PYRITINE MET 0
SQ SEQUENCE 399 AA; 45766 MW; 37F0C0D2C3BA5C45 CRC64;

Query Match 67.8%; Score 40; DB 1; Length 399;
Best Local Similarity 85.7%; Pred. No. 41;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDD 8
DB 38 FRCRGDD 44
:|||||

RESULT 14
PNCB_SALTI STANDARD; PRT; 399 AA.
AC Q827Y9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Nicotinate phosphoribosyltransferase (EC 2.4.2.11) (NAPRTase).
GN Name=pncB; OrderedLocusNames=ST1010, t1930;
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;

```

```

RX MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Felwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RX DOI=10.1128/JB.185.7.2330-2337.2003;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18.";
RL J. Bacteriol. 185:2330-2337(2003).
CC -1- CATALYTIC ACTIVITY: Nicotinate D-ribonucleotide + diphosphate =
CC nicotinate + 5-phospho-alpha-D-ribose 1-diphosphate.
CC -1- PATHWAY: NAD biosynthesis; nicotinamide to NAMN; second step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the NAPRTase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL627268; CAD05404.1; --
CC EMBL; AE018840; AAO69546.1; --
CC HAMAP; MF_00570; --; 1.
CC InterPro; IPR007229; NAPRTase.
CC InterPro; IPR006406; Nic Prtrans.
CC InterPro; IPR008967; P53_like_DNA_bnd.
CC Pfam; PF04095; NAPRTase; 1.
CC PIRSF; PIRSF000484; Nicot_phos_ribo; 1.
CC TIGRFAMs; TIGR01514; NAPRTase; 1.
CC Complete proteome; Glycosyltransferase;
KW Pyridine nucleotide biosynthesis; Transferase.
FT INIT MET 0
FT PYRITINE MET 0
SQ SEQUENCE 399 AA; 45562 MW; ADF38EAE4618214 CRC64;

Query Match 67.8%; Score 40; DB 1; Length 399;
Best Local Similarity 85.7%; Pred. No. 41;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDD 8
DB 38 FRCRGDD 44
:|||||

RESULT 15
PNCB_SALTY STANDARD; PRT; 399 AA.
AC P22253;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Nicotinate phosphoribosyltransferase (EC 2.4.2.11) (NAPRTase).
GN Name=pncB; OrderedLocusNames=STM1004;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;

```

RX	MEDLINE=22590274; PubMed=12704152;
RA	DOI=10.1128/JAI.71.5.2775-2786.2003;
RR	Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA	Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA	Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA	Schwartz D.C., Blattner F.R.;
RT	"Complete genome sequence and comparative genomics of Shigella
RT	flexneri serotype 2a strain 245T7.";
RL	Infect. Immun. 71:2775-2786(2003).
DR	EMBL; AS016981; AAP16443.1; -.
DR	GO; GO:0004516; F.nicotinate phosphoribosyltransferase activity; IEA.
DR	GO; GO:0016757; F.transferase activity, transferring glycosyl... ; IEA.
DR	GO; GO:0019357; F.nicotinate nucleotide biosynthesis; IEA.
DR	GO; GO:0019363; P.pyridine nucleotide biosynthesis; IEA.
DR	InterPro; IPR007229; NAPRTase.
DR	InterPro; IPR006406; Nic_Prttrans.
DR	InterPro; IPR008967; P53_like_DNA_bnd.
DR	Pfam; PF04095; NAPRTase; 1.
DR	TIGRFAM; TIGR00484; Nicot_phos_ribo; 1.
DR	TIGRFAM; TIGR01514; NAPRTase; 1.
KW	Glycosyltransferase; Transferase.
SQ	SEQUENCE 400 AA; 45911 MW; 5385BCB1A8703095 CRC64;
QY	Query Match 67.8%; Score 40; DB 2; Length 400;
DB	Best Local Similarity 85.7%; Pred.No. 41;
	Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps
RESULT 17	
Q83LN3	
ID	Q83LN3 PRELIMINARY; PRT; 416 AA.
AC	Q83LN3;
DT	01-JUN-2003 (TREMBLrel. 24, Created)
DT	01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE	Nicotinate phosphoribosyltransferase.
GN	Name=pncB; OrderedLocusNames=SF0928;
OS	Shigella flexneri.
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC	Enterobacteriaceae; Shigella.
OX	NCBI_TaxID=623;
RN	[1] _
RP	SEQUENCE FROM N.A.
RC	STRAIN=301 / Serotype 2a;
RX	MEDLINE=2272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA	Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA	Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA	Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA	Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA	Yu J.;
RT	"Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT	through comparison with genomes of Escherichia coli K12 and O157.";
RL	Nucleic Acids Res. 30:4432-4441(2002).
DR	EMBL; AS015122; AAN42557.1; -.
DR	GO; GO:0004516; F.nicotinate phosphoribosyltransferase activity; IEA.
DR	GO; GO:0016757; F.transferase activity, transferring glycosyl... ; IEA.
DR	GO; GO:0019357; F.nicotinate nucleotide biosynthesis; IEA.
DR	GO; GO:0019363; P.pyridine nucleotide biosynthesis; IEA.
DR	InterPro; IPR007229; NAPRTase.
DR	InterPro; IPR006406; Nic_Prttrans.
DR	InterPro; IPR008967; P53_like_DNA_bnd.
DR	Pfam; PF04095; NAPRTase; 1.
DR	PIRS; PIRSF000484; Nicot_phos_ribo; 1.
DR	TIGRFAM; TIGR01514; NAPRTase; 1.
KW	Complete proteome; Glycosyltransferase; Transferase.
SQ	SEQUENCE 416 AA; 47700 MW; E43D723542AF2DD5 CR664;
QY	Query Match 67.8%; Score 40; DB 2; Length 416;
DB	Best Local Similarity 85.7%; Pred.No. 42;

```

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDD 8
Db 55 FRCRGDD 61

RESULT 18
Q8FJ98 PRELIMINARY; PRT; 416 AA.
AC Q8FJ98;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nicotinate phosphoribosyltransferase (EC 2.4.2.11).
GN Name=ncn8; OrderedLocusNames=ci073;
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O6:H1 / CFT073 / ATCC 700928 / UPEC;
RX MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
DR EMBL; AE016758; AN79541.1; -.
DR GO; GO:0004516; P:nicotinate phosphoribosyltransferase activity; IEA.
DR GO; GO:0019357; P:nicotinate nucleotide biosynthesis; IEA.
DR GO; GO:0019363; P:pyridine nucleotide biosynthesis; IEA.
DR InterPro; IPR007229; NAPTase.
DR InterPro; IPR006406; Nic_Prtase.
DR InterPro; IPR008967; P53_like_DNA_bnd.
DR Pfam; PF04095; NAPTase; 1.
DR TIGSF; TIGSF00484; Nicot_phos_ribo; 1.
DR TIGRFAMs; TIGR01514; NAPTase; 1.
KW Complete proteome.
SQ SEQUENCE 416 AA; 47722 MW; F00AFC07B04B1E23 CRC64;

Query Match 67.8%; Score 40; DB 2; Length 416;
Best Local Similarity 85.7%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDD 8
Db 55 FRCRGDD 61

RESULT 19
Q50468 PRELIMINARY; PRT; 705 AA.
AC Q50468;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Pks002a.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RA Smith D.R.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP Robison K.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.

```

```

CC -!- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
CC family.
EMBL; U00024; AAA50930.1; -.
DR HSSP; P08659; 1LCI.
DR GO; GO:0003824; P:catalytic activity; IEA.
DR GO; GO:0048037; P:cofactor binding; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR009081; ACP_like.
DR InterPro; IPR000873; AMP-bind.
DR InterPro; IPR006163; Sheppanteth_bind.
DR Pfam; PF00501; PP-binding; 1.
DR Pfam; PF00550; PP-binding; 1.
DR PRINTS; PR00154; AMPEBINDING.
DR PROSITE; PS50075; ACP_DOMAIN; 1.
SQ SEQUENCE 705 AA; 75169 MW; 9F9D4F500690BF5D CRC64;

Query Match 67.8%; Score 40; DB 2; Length 705;
Best Local Similarity 66.7%; Pred. No. 73;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 YRCRGDDSK 10
Db 383 YRCRADDT 391

RESULT 20
P96283 PRELIMINARY; PRT; 705 AA.
AC P96283; Q7D6D9;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE PROBABLE FATTY-ACID-CoA LIGASE FADD22 (FATTY-ACID-CoA SYNTHETASE)
DE (FATTY-ACID-CoA SYNTHASE) (EC 6.2.1.1-) (Substrate--CoA ligase).
GN Name=fadd22; OrderedLocusNames=MT3021, RV2948c;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eigmeier K., Gas S., Barry C.E. III,
RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA Holtroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulten J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490 (2002).
CC -!- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
CC family.
DR EMBL; BX842581; CAB06101.1; -.
DR EMBL; AE000516; AAK47347.1; -.
DR PIR; A70669; A70669.

```

```
DR HSSP; P08659; 1LCI.
DR TIGR; MT3021; -.
DR Tuberculist; Bv2948c; -.
DR GO; GO:0003824; F: catalytic activity; IEA.
DR GO; GO:004803; F: cofactor binding; IEA.
DR GO; GO:0016874; F: ligase activity; IEA.
DR GO; GO:0008152; P: metabolism; IEA.
DR InterPro; IPR009081; ACP like.
DR InterPro; IPR000873; AMP-bind.
DR InterPro; IPR006163; Phosphanteth_bind.
DR Pfam; PF00501; AMP-binding; 1.
DR Pfam; PF00550; PP-binding; 1.
DR PRINTS; PRO0154; AMPBINDING.
DR PROSITE; PS50075; ACP_DOMAIN; 1.
KW Ligase.
SQ SEQUENCE 705 AA; 75197 MW; 600F2D0EABFDF1DC CRC64;

Query Match 67.8%; Score 40; DB 2; Length 705;
Best Local Similarity 66.7%; Pred. No. 73;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 YRCRGDSDK 10
DB |||||:::
383 YRCRADDTT 391

RESULT 21
Q7TXK7
ID Q7TXK7 PRELIMINARY; PRT; 705 AA.
AC Q7TXK7
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-NAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PROBABLE FATTY-ACID-CoA LIGASE FAD22 (FATTY-ACID-CoA SYNTHETASE)
DE (FATTY-ACID-CoA SYNTHASE) (EC 6.2.1.1.-)
DE Name=fad22; OrderedLocusNames=Mb2972c;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrall B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR EMBL; BX248344; CAD96659.1; -.
DR GO; GO:0048037; F: cofactor binding; IEA.
DR GO; GO:0016874; F: ligase activity; IEA.
DR GO; GO:0008152; P: metabolism; IEA.
DR InterPro; IPR009081; ACP like.
DR InterPro; IPR000873; AMP-bind.
DR InterPro; IPR006163; Phosphanteth_bind.
DR Pfam; PF00501; AMP-binding; 1.
DR Pfam; PF00550; PP-binding; 1.
DR PROSITE; PS50075; ACP_DOMAIN; 1.
KW Complete proteome; Ligase.
SQ SEQUENCE 705 AA; 75197 MW; 600F2D0EABFDF1DC CRC64;

Query Match 67.8%; Score 40; DB 2; Length 705;
Best Local Similarity 66.7%; Pred. No. 73;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 YRCRGDSDK 10
DB |||||:::
383 YRCRADDTT 391

RESULT 22
Q7TXK7
ID Q7TXK7 PRELIMINARY; PRT; 705 AA.
AC Q7TXK7
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-NAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PROBABLE FATTY-ACID-CoA LIGASE FAD22 (FATTY-ACID-CoA SYNTHETASE)
DE (FATTY-ACID-CoA SYNTHASE) (EC 6.2.1.1.-)
DE Name=fad22; OrderedLocusNames=Mb2972c;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrall B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR EMBL; BX248344; CAD96659.1; -.
DR GO; GO:0048037; F: cofactor binding; IEA.
DR GO; GO:0016874; F: ligase activity; IEA.
DR GO; GO:0008152; P: metabolism; IEA.
DR InterPro; IPR009081; ACP like.
DR InterPro; IPR000873; AMP-bind.
DR InterPro; IPR006163; Phosphanteth_bind.
DR Pfam; PF00501; AMP-binding; 1.
DR Pfam; PF00550; PP-binding; 1.
DR PROSITE; PS50075; ACP_DOMAIN; 1.
KW Complete proteome; Ligase.
SQ SEQUENCE 705 AA; 75197 MW; 600F2D0EABFDF1DC CRC64;

Query Match 67.8%; Score 40; DB 2; Length 705;
Best Local Similarity 66.7%; Pred. No. 73;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 YRCRGDSDK 10
DB |||||:::
383 YRCRADDTT 391
```

```
Q8WQJ3
ID Q8WQJ3 PRELIMINARY; PRT; 845 AA.
AC Q8WQJ3
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE SRCR LCCL adhesive-like protein (Fragment).
GN Name=slap;
OS Plasmodium chabaudi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5825;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21982590; PubMed=11985859; DOI=10.1016/S0166-6851(02)00016-6;
RA Delrieu I., Waller C.C., Mota M.M., Grainger M., Langhorne J.,
RA Holder A.A.;
RT "PSLAP, a protein with multiple adhesive motifs, is expressed in
RT Plasmodium falciparum gametocytes.";
RL Mol. Biochem. Parasitol. 121:11-20(2002).
DR EMBL; AY072023; AAL58521.1; -.
DR GO; GO:0016020; C: membrane; IEA.
DR GO; GO:0005044; F: scavenger receptor activity; IEA.
DR Pfam; PF03815; LCCL; 3.
DR Pfam; PF00530; SRCR; 2.
DR SMART; SM00202; SR; 2.
DR PROSITE; PS50820; LCCL; 2.
DR PROSITE; PS50287; SRCR_2; 2.
FT NON_TER 1
FT NON_TER 845
SQ SEQUENCE 845 AA; 95032 MW; 06A389E4CC294477 CRC64;

Query Match 67.8%; Score 40; DB 2; Length 845;
Best Local Similarity 55.6%; Pred. No. 88;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYRCRGDSDS 9
DB ||::|:::
340 NFRCKGDEA 348

RESULT 23
Q8WT63
ID Q8WT63 PRELIMINARY; PRT; 1304 AA.
AC Q8WT63
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Multidomain scavenger receptor protein pBSR precursor.
OS Plasmodium berghei.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5821;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22242506; PubMed=12354219;
RA Claudianos C., Dessens J.T., Trueman H.F., Arai M., Mendoza J.,
RA Butcher G.A., Crompton T., Sinden R.E.;
RT "A malaria scavenger receptor-like protein essential for parasite
RT development.";
RL Mol. Microbiol. 45:1473-1484(2002).
CC -!- SIMILARITY: Contains 1 PLAT domain.
DR EMBL; AY034780; AAK64185.1; -.
DR HSSP; O43405; 1UJBI.
DR GO; GO:0016020; C: membrane; IEA.
DR GO; GO:0004872; F: receptor activity; IEA.
DR GO; GO:0005044; F: scavenger receptor activity; IEA.
DR Pfam; PF03815; LCCL; 4.
DR Pfam; PF01477; PLAT; 1.
DR Pfam; PF00530; SRCR; 2.
DR SMART; SM00202; SR; 2.
DR PROSITE; PS50820; LCCL; 3.
DR PROSITE; PS50095; PLAT; 1.
DR PROSITE; PS50287; SRCR_2; 2.
KW Receptor; Signal.
```

```

FT SIGNAL 1 22 Potential.
SQ SEQUENCE 1304 AA; 148247 MW; F936CC94E7B19F1C CRC64;

Query Match 67.8%; Score 40; DB 2; Length 1304;
Best Local Similarity 55.6%; Pred. No. 1.4e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYRCRGDGS 9
   |||:|:|:|:|
Db 614 NFRCKGDEA 622

RESULT 24
Q8P0M2 PRELIMINARY; PRT; 37 AA.
ID Q8P0M2 AC Q8P0M2 PRELIMINARY; PRT; 154 AA.
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical phase protein.
DE OrderedLocuNames=spym18_1296;
GN Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=186103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS8232;
RX MEDLINE=21927593; PubMed=11917108; DOI=10.1073/pnas.062526099;
RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
DR EMBL; AB010051; ALU97902.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 37 AA; 4398 MW; E7884FEAA642342 CRC64;

Query Match 66.1%; Score 39; DB 2; Length 37;
Best Local Similarity 70.0%; Pred. No. 5.4;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NYRCRGDGS 10
   |||:|:|:|
Db 14 NYHTNGDGS 23

RESULT 25
Q848W4 PRELIMINARY; PRT; 46 AA.
ID Q848W4 AC Q848W4 PRELIMINARY; PRT; 154 AA.
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
DE Bacillus megaterium.
OS Plasmid pBM400.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1404;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=QM B1551;
RX MEDLINE=22964369; PubMed=14602653;
RX DOI=10.1128/AEM.69.11.6888-6898.2003;
RA Scholte M.D., White C.A., Kummalaiaayan M., Vary P.S.;
RT "Sequencing and characterization of pBM400 from Bacillus megaterium QM
RT B1551.";
RL Appl. Environ. Microbiol. 69:6888-6898(2003).
DR EMBL; AF142677; AA052802.1; -.
KW Hypothetical protein; Plasmid.

```

```

SQ SEQUENCE 46 AA; 5449 MW; F4E4C4A3A050E138 CRC64;

Query Match 66.1%; Score 39; DB 2; Length 46;
Best Local Similarity 60.0%; Pred. No. 6.8;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYRCRGDGS 10
   |||:|:|:|
Db 35 NYRCRGEDVK 44

RESULT 26
Q8BUS2 PRELIMINARY; PRT; 154 AA.
ID Q8BUS2 AC Q8BUS2 PRELIMINARY; PRT; 154 AA.
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Mus musculus ES cells cDNA, RIKEN full-length enriched library,
DE clone:C330016E03 product:SRCRB-S4D PROTEIN homolog (Fragment).
GN Name=Srcrbad;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata K., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsumi T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;

```

```
RA Adachi J., Aizawa K., Akahira S., Akimura T., Aono H., Arai A.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Iehi Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yamanaka I., Yasunishi A.,
RA Yoshida K., Yoshino M., Muramatsu M., Hayashizaki Y.,
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ Databases.
RL EMBL; AK082786; BAC38619.1; -.
DR HSSP; Q08380; 1BY2.
DR MGD; MGI:1924709; Srcrb4d.
DR GO; GO:0016020; C-membrane; IEA.
DR GO; GO:0005044; F-scravenger receptor activity; IEA.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00530; SRCR; 1.
DR PRINTS; PR00258; SPERACTRCPTR.
DR SMART; SM00202; SR; 1.
DR PROSITE; PS00420; SRCR_1; 1.
DR PROSITE; PS0287; SRCR_2; 1.
FT NON_TER 1
SQ SEQUENCE 154 AA; 17028 MW; 8932AA223E2F14F5 CRC64;

Query Match 66.1%; Score 39; DB 2; Length 154;
Best Local Similarity 66.7%; Pred. No. 24;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NVRCRGDS 9
DB 118 NVKRGDES 126
|:|||||
|:|||||

RESULT 27
Q20683
ID Q20683 PRELIMINARY; PRT; 179 AA.
AC Q20683;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein F52H3.5.
GN ORFNames=F52H3.5;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN SEQUENCE FROM N.A.
RP STRAIN=Bristol N2;
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Gardner A.E.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z66512; CAA91325.1; -.
DR PIR; T22521; T22521.
DR WormBase; WBGene0009947; F52H3.5.
DR WormPep; F52H3.5; CE03401.
DR InterPro; IPR001440; TPR.
DR InterPro; IPR008941; TPR-like.
DR Pfam; PF00515; TPR_1; 1.
DR SMART; SM00028; TPR; 3.
DR PROSITE; PS50005; TPR; 1.
DR PROSITE; PS50293; TPR_REGION; 1.
KW Hypothetical protein_Repeat; TPR repeat.
SQ SEQUENCE 179 AA; 19648 MW; 6FF899837C65A18F CRC64;
```

```
Query Match 66.1%; Score 39; DB 2; Length 179;
Best Local Similarity 77.8%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 YRCRGDDSK 10
DB 125 YRLRGDDDK 133
|:|||||
|:|||||

RESULT 28
O70681 PRELIMINARY; PRT; 588 AA.
AC O70681;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Coat protein (Fragment).
OS Sugarcane streak mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyvirus.
OX NCBI_TaxID=53954;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99162284; PubMed=10051385; DOI=10.1006/mpev.1998.0535;
RA Hall J.S., Adams B., Parsons T.J., French R., Lane L.C., Jensen S.G.;
RT "Molecular cloning, sequencing, and phylogenetic relationships of a
RT new potyvirus: sugarcane streak mosaic virus, and a reevaluation of
RT the classification of the potyviridae.";
RL Mol. Phylogenet. Evol. 10:323-332(1998).
CC -!- SIMILARITY: Belongs to the potyviruses polyprotein family.
DR EMBL; U75456; AAC16271.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0019079; F:viral genome replication; IEA.
DR InterPro; IPR001592; Poty_coat.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_P5vir.
DR Pfam; PF00767; Poty_coat; 1.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
KW Coat protein.
FT NON_TER 1
SQ SEQUENCE 588 AA; 66335 MW; 3876EEF59189EA32 CRC64;

Query Match 66.1%; Score 39; DB 2; Length 588;
Best Local Similarity 75.0%; Pred. No. 93;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NVRCRGDD 8
DB 141 NYTCRGDD 148
|:|||||
|:|||||

RESULT 29
Q3YNE1 PRELIMINARY; PRT; 1420 AA.
ID Q3YNE1;
AC Q3YNE1;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein (Fragment).
OS Sugarcane streak mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyvirus.
OX NCBI_TaxID=53954;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99243201; PubMed=10226614;
RA Hema M., Joseph J., Gopinath K., Sreenivasulu P., Savithri H.S.;
```

RT "Molecular characterization and intervirial relationships of a flexuous
 RT filamentous virus causing mosaic disease of sugarcane (Saccharum
 RT officinarum L.) in India."
 RL Arch. Virol. 144:479-490(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Hema M.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the potyviruses polyprotein family.
 DR EMBL; Y17738; CAA76842.3; -.
 DR GO: GO:0019028; C: viral capsid; IEA.
 DR GO: GO:0008234; F: cysteine-type peptidase activity; IEA.
 DR GO: GO:0003723; F: RNA binding; IEA.
 DR GO: GO:0003723; F: RNA binding; IEA.
 DR GO: GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
 DR GO: GO:0005198; F: structural molecule activity; IEA.
 DR GO: GO:0006508; P: proteolysis and peptidolysis; IEA.
 DR GO: GO:0006350; P: transcription; IEA.
 DR GO: GO:0019079; P: viral genome replication; IEA.
 DR InterPro; IPR001730; Peptidase_C4.
 DR InterPro; IPR001592; Poty_coat.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR001205; RNA_pol_P3B.
 DR InterPro; IPR007094; RNA_pol_P5vir.
 DR Pfam; PF00767; Poty_coat_1.
 DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
 DR PRINTS; PR00966; NIAPOTYPASE.
 DR Coar protein; Polypoteine.
 FT NON TER 1 1
 FT CHAIN 1138 1420 coat protein.
 SQ SEQUENCE 1420 AA; 160420 MW; 488B456258D87275 CRC64;
 Query Match 66.1%; Score 39; DB 2; Length 1420;
 Best Local Similarity 75.0%; Pred. No. 2.3e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 NYRCRGDD 8
 DB 974 NYTCNGDD 981
 RESULT 30
 Q7RQM4 PRELIMINARY; PRT; 1615 AA.
 AC Q7RQM4
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Multidomain scavenger receptor protein PbsR precursor.
 GN Name=PY01071;
 OS Plasmodium yoelii yoelii.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 ON NCBI_TaxID=73239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=17XNU;
 RX PubMed=12368865; DOI=10.1038/nature01099;
 RA Carlton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Perteau M.,
 RA Silva J.C., Ermolaeva M.D., Allen J.E., Selegut J.D., Koo H.L.,
 RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
 RA Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
 RA Cho J.K., Quackenbush J., Sedegh M., Shoaibi A., Cummings L.M.,
 RA Florens L., Yates P.R. III, Raine J.D., Sinden R.E., Harris M.A.,
 RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
 RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
 RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
 RA Carucci D.J.;
 RT "Genome sequence and comparative analysis of the model rodent malaria
 RT parasite Plasmodium yoelii yoelii."
 RL Nature 419:512-519(2002).
 CC -1- SIMILARITY: Contains 1 PLAT domain.
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.

DR EMBL; ABL01000281; EAA20072.1; -.
 DR HSSP; O57811; IG8A.
 DR GO: GO:0016020; C: membrane; IEA.
 DR GO: GO:0005634; C: nucleus; IEA.
 DR GO: GO:0004872; F: receptor activity; IEA.
 DR GO: GO:0003723; F: RNA binding; IEA.
 DR GO: GO:0005044; F: scavenger receptor activity; IEA.
 DR GO: GO:0006364; F: RNA processing; IEA.
 DR InterPro; IPR000895; ConA_like lec_gl.
 DR InterPro; IPR000692; Fibrillarin.
 DR InterPro; IPR004043; LCCL.
 DR InterPro; IPR001024; Lipoxigenase_LH2.
 DR InterPro; IPR008976; PLAT_LH2.
 DR InterPro; IPR001190; Srcr_receptor.
 DR Pfam; PF01269; Fibrillarin; 1.
 DR Pfam; PF03815; LCCL; 4.
 DR Pfam; PF01477; PLAT; 1.
 DR Pfam; PF00530; SRCR; 2.
 DR PRINTS; PR00052; FIBRILLARIN.
 DR ProDom; PD004637; Fibrillarin; 1.
 DR PROSITE; PS00566; FIBRILLARIN; 1.
 DR PROSITE; PS00820; LCCL; 4.
 DR PROSITE; PS00095; PLAT; 1.
 DR PROSITE; PS00287; SRCR_2; 2.
 KW Receptor.
 SQ SEQUENCE 1615 AA; 180943 MW; 94053C4C9AC93505 CRC64;
 Query Match 66.1%; Score 39; DB 2; Length 1615;
 Best Local Similarity 62.5%; Pred. No. 2.6e+02;
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NYRCRGDD 8
 DB 614 NFRCKGDE 621
 RESULT 31
 Q8VI63 PRELIMINARY; PRT; 121 AA.
 AC Q8VI63
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Lambda 3 protein (fragment).
 OS Reovirus sp.
 OC Viruses; dsRNA viruses; Reoviridae; unclassified Reoviridae.
 ON NCBI_TaxID=10891;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21920879; PubMed=11923358;
 RA Leary T.P., Erker J.C., Chalmers M.L., Cruz A.T., Wetzel J.D.,
 RA Desai S.M., Mushahwar I.K., Dermody T.S.;
 RT "Detection of mammalian reovirus RNA by using reverse transcription-
 RT PCR: sequence diversity within the lambda3-encoding L1 gene."
 RL J. Clin. Microbiol. 40:1368-1375(2002).
 DR EMBL; AY007423; AAG28882.1; -.
 DR HSSP; P17378; IN35.
 DR GO: GO:0003723; F: RNA binding; IEA.
 DR GO: GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
 DR GO: GO:0019079; P: viral genome replication; IEA.
 FT NON TER 1 121
 FT NON TER 121 121
 SQ SEQUENCE 121 AA; 13363 MW; EA44987394A246CA CRC64;
 Query Match 64.4%; Score 38; DB 2; Length 121;
 Best Local Similarity 75.0%; Pred. No. 28;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NYRCRGDD 8
 DB 96 NYVCQGGD 103

RESULT 32

Q8V164
ID Q8V164 PRELIMINARY; PRT; 121 AA.
AC Q8V164;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Lambda 3 protein (Fragment).
OS Reovirus sp.
OC Viruses; dsRNA viruses; Reoviridae; unclassified Reoviridae.
OX NCBI_TaxID=10891;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21920879; PubMed=11923358;
RA Leary T.P., Erker J.C., Chalmers M.L., Cruz A.T., Wetzel J.D.,
Desai S.M., Mushahwar I.K., Dermody T.S.;
RT "Detection of mammalian reovirus RNA by using reverse transcription-
PCR: sequence diversity within the lambda3-encoding L1 gene.";
RL J. Clin. Microbiol. 40:1368-1375(2002).
DR EMBL; AY007422; AAG28881.1; -.
DR HSSP; P17378; IN35.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
FT NON_TER 1
FT NON_TER 121
SQ SEQUENCE 121 AA; 13349 MW; 2B8DA9F9BC00CCFB CRC64;

Query Match 64.4%; Score 38; DB 2; Length 121;
Best Local Similarity 75.0%; Pred. No. 28;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDD 8
|||:||||
Db 96 NYVCGDD 103

RESULT 33

Q8V165
ID Q8V165 PRELIMINARY; PRT; 121 AA.
AC Q8V165;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Lambda 3 protein (Fragment).
OS Reovirus sp.
OC Viruses; dsRNA viruses; Reoviridae; unclassified Reoviridae.
OX NCBI_TaxID=10891;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21920879; PubMed=11923358;
RA Leary T.P., Erker J.C., Chalmers M.L., Cruz A.T., Wetzel J.D.,
Desai S.M., Mushahwar I.K., Dermody T.S.;
RT "Detection of mammalian reovirus RNA by using reverse transcription-
PCR: sequence diversity within the lambda3-encoding L1 gene.";
RL J. Clin. Microbiol. 40:1368-1375(2002).
DR EMBL; AY007421; AAG28880.1; -.
DR HSSP; P17378; IN35.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
FT NON_TER 1
FT NON_TER 121
SQ SEQUENCE 121 AA; 13363 MW; EA44987394A246CA CRC64;

Query Match 64.4%; Score 38; DB 2; Length 121;
Best Local Similarity 75.0%; Pred. No. 28;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDD 8
|||:||||
Db 96 NYVCGDD 103

RESULT 34

Q8V166
ID Q8V166 PRELIMINARY; PRT; 121 AA.
AC Q8V166;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Lambda 3 protein (Fragment).
OS Reovirus sp.
OC Viruses; dsRNA viruses; Reoviridae; unclassified Reoviridae.
OX NCBI_TaxID=10891;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21920879; PubMed=11923358;
RA Leary T.P., Erker J.C., Chalmers M.L., Cruz A.T., Wetzel J.D.,
Desai S.M., Mushahwar I.K., Dermody T.S.;
RT "Detection of mammalian reovirus RNA by using reverse transcription-
PCR: sequence diversity within the lambda3-encoding L1 gene.";
RL J. Clin. Microbiol. 40:1368-1375(2002).
DR EMBL; AY007420; AAG28879.1; -.
DR HSSP; P17378; IN35.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
FT NON_TER 1
FT NON_TER 121
SQ SEQUENCE 121 AA; 13363 MW; EA44987394A246CA CRC64;

Query Match 64.4%; Score 38; DB 2; Length 121;
Best Local Similarity 75.0%; Pred. No. 28;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDD 8
|||:||||
Db 96 NYVCGDD 103

RESULT 35

Q8V167
ID Q8V167 PRELIMINARY; PRT; 121 AA.
AC Q8V167;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Lambda 3 protein (Fragment).
OS Reovirus sp.
OC Viruses; dsRNA viruses; Reoviridae; unclassified Reoviridae.
OX NCBI_TaxID=10891;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21920879; PubMed=11923358;
RA Leary T.P., Erker J.C., Chalmers M.L., Cruz A.T., Wetzel J.D.,
Desai S.M., Mushahwar I.K., Dermody T.S.;
RT "Detection of mammalian reovirus RNA by using reverse transcription-
PCR: sequence diversity within the lambda3-encoding L1 gene.";
RL J. Clin. Microbiol. 40:1368-1375(2002).
DR EMBL; AY007419; AAG28878.1; -.
DR HSSP; P17378; IN35.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
FT NON_TER 1
FT NON_TER 121
SQ SEQUENCE 121 AA; 13363 MW; EA44987394A246CA CRC64;

Query Match 64.4%; Score 38; DB 2; Length 121;
Best Local Similarity 75.0%; Pred. No. 28;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDD 8
|||:||||
Db 96 NYVCGDD 103


```
Db          96 NYVCGDD 103

RESULT 36
Q8V168      PRELIMINARY;      PRT;      121 AA.
AC Q8V168;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Lambda 3 protein (Fragment).
OS Reovirus sp.
OC Viruses; dsRNA viruses; Reoviridae; unclassified Reoviridae.
OX NCBI_TaxID=10891;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21920879; PubMed=11923358;
RA Leary T.P., Erker J.C., Chalmers M.L., Cruz A.T., Wetzel J.D.,
RA Desai S.M., Mushahwar I.K., Dermody T.S.;
RT "Detection of mammalian reovirus RNA by using reverse transcription-
RT PCR: sequence diversity within the lambda3-encoding L1 gene.";
RL J. Clin. Microbiol. 40:1368-1375(2002).
DR EMBL; AY007418; AAG28877.1; -.
DR HSSP; P17378; IN35.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0019079; P:Viral genome replication; IEA.
FT NON_TER 1
FT NON_TER 121
SQ SEQUENCE 121 AA; 13363 MW; EA44987394A246CA CRC64;

Query Match      64.4%; Score 38; DB 2; Length 121;
Best Local Similarity 75.0%; Pred. No. 28;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDD 8
||| :|||
Db 96 NYVCGDD 103

RESULT 37
Q8V169      PRELIMINARY;      PRT;      121 AA.
AC Q8V169;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Lambda 3 protein (Fragment).
OS Reovirus sp.
OC Viruses; dsRNA viruses; Reoviridae; unclassified Reoviridae.
OX NCBI_TaxID=10891;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21920879; PubMed=11923358;
RA Leary T.P., Erker J.C., Chalmers M.L., Cruz A.T., Wetzel J.D.,
RA Desai S.M., Mushahwar I.K., Dermody T.S.;
RT "Detection of mammalian reovirus RNA by using reverse transcription-
RT PCR: sequence diversity within the lambda3-encoding L1 gene.";
RL J. Clin. Microbiol. 40:1368-1375(2002).
DR EMBL; AY007417; AAG28876.1; -.
DR HSSP; P17378; IN35.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0019079; P:Viral genome replication; IEA.
FT NON_TER 1
FT NON_TER 121
SQ SEQUENCE 121 AA; 13363 MW; EA44987394A246CA CRC64;

Query Match      64.4%; Score 38; DB 2; Length 121;
Best Local Similarity 75.0%; Pred. No. 28;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDD 8
||| :|||
Db 96 NYVCGDD 103

RESULT 38
Q8V170      PRELIMINARY;      PRT;      121 AA.
AC Q8V170;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Lambda 3 protein (Fragment).
OS Reovirus sp.
OC Viruses; dsRNA viruses; Reoviridae; unclassified Reoviridae.
OX NCBI_TaxID=10891;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21920879; PubMed=11923358;
RA Leary T.P., Erker J.C., Chalmers M.L., Cruz A.T., Wetzel J.D.,
RA Desai S.M., Mushahwar I.K., Dermody T.S.;
RT "Detection of mammalian reovirus RNA by using reverse transcription-
RT PCR: sequence diversity within the lambda3-encoding L1 gene.";
RL J. Clin. Microbiol. 40:1368-1375(2002).
DR EMBL; AY007416; AAG28875.1; -.
DR HSSP; P17378; IN35.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0019079; P:Viral genome replication; IEA.
FT NON_TER 1
FT NON_TER 121
SQ SEQUENCE 121 AA; 13363 MW; EA44987394A246CA CRC64;

Query Match      64.4%; Score 38; DB 2; Length 121;
Best Local Similarity 75.0%; Pred. No. 28;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDD 8
||| :|||
Db 96 NYVCGDD 103

RESULT 39
Q8V171      PRELIMINARY;      PRT;      121 AA.
AC Q8V171;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Lambda 3 protein (Fragment).
OS Reovirus sp.
OC Viruses; dsRNA viruses; Reoviridae; unclassified Reoviridae.
OX NCBI_TaxID=10891;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21920879; PubMed=11923358;
RA Leary T.P., Erker J.C., Chalmers M.L., Cruz A.T., Wetzel J.D.,
RA Desai S.M., Mushahwar I.K., Dermody T.S.;
RT "Detection of mammalian reovirus RNA by using reverse transcription-
RT PCR: sequence diversity within the lambda3-encoding L1 gene.";
RL J. Clin. Microbiol. 40:1368-1375(2002).
DR EMBL; AY007415; AAG28874.1; -.
DR HSSP; P17378; IN35.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0019079; P:Viral genome replication; IEA.
FT NON_TER 1
FT NON_TER 121
SQ SEQUENCE 121 AA; 13363 MW; EA44987394A246CA CRC64;

Query Match      64.4%; Score 38; DB 2; Length 121;
Best Local Similarity 75.0%; Pred. No. 28;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDD 8
||| :|||
Db 96 NYVCGDD 103
```

QY 1 NYRCRGDD 8

Db 96 NYVCGDD 103
|| |:|||

RESULT 40

Q8V172 PRELIMINARY; PRT; 121 AA.
AC Q8V172;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Lambda 3 protein (fragment).
OS Reovirus sp.
OC Viruses; dsRNA viruses; Reoviridae; unclassified Reoviridae.
OX NCBI_TaxID=10891;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21920879; PubMed=11923358;
RA Leary T.P., Erker J.C., Chalmers M.L., Cruz A.T., Wetzel J.D.,
RA Desai S.M., Mushahwar I.K., Dermody T.S.;
RT "Detection of mammalian reovirus RNA by using reverse transcription-PCR: sequence diversity within the lambda3-encoding L1 gene.";
RT J. Clin. Microbiol. 40:1368-1375(2002).
RL EMBL: AY007414; AAG28872.1; -.
DR HSSP; P17378; 1N35.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0019079; P:Viral genome replication; IEA.
FT NON_TER 1
FT NON_TER 121
SQ SEQUENCE 121 AA; 13391 MW; 4A53987DB0292B7C CRC64;

Query Match 64.4%; Score 38; DB 2; Length 121;
Best Local Similarity 75.0%; Pred. No. 28;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDD 8
|| |:|||

Db 96 NYVCGDD 103

RESULT 41

Q8V173 PRELIMINARY; PRT; 121 AA.
AC Q8V173;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Lambda 3 protein (fragment).
OS Reovirus sp.
OC Viruses; dsRNA viruses; Reoviridae; unclassified Reoviridae.
OX NCBI_TaxID=10891;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21920879; PubMed=11923358;
RA Leary T.P., Erker J.C., Chalmers M.L., Cruz A.T., Wetzel J.D.,
RA Desai S.M., Mushahwar I.K., Dermody T.S.;
RT "Detection of mammalian reovirus RNA by using reverse transcription-PCR: sequence diversity within the lambda3-encoding L1 gene.";
RT J. Clin. Microbiol. 40:1368-1375(2002).
RL EMBL: AY007413; AAG28872.1; -.
DR HSSP; P17378; 1N35.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0019079; P:Viral genome replication; IEA.
FT NON_TER 1
FT NON_TER 121
SQ SEQUENCE 121 AA; 13391 MW; 4A53987DB0292B7C CRC64;

Query Match 64.4%; Score 38; DB 2; Length 121;
Best Local Similarity 75.0%; Pred. No. 28;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDD 8
|| |:|||

Db 96 NYVCGDD 103

RESULT 42

Q8V174 PRELIMINARY; PRT; 121 AA.
AC Q8V174;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Lambda 3 protein (fragment).
OS Reovirus sp.
OC Viruses; dsRNA viruses; Reoviridae; unclassified Reoviridae.
OX NCBI_TaxID=10891;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21920879; PubMed=11923358;
RA Leary T.P., Erker J.C., Chalmers M.L., Cruz A.T., Wetzel J.D.,
RA Desai S.M., Mushahwar I.K., Dermody T.S.;
RT "Detection of mammalian reovirus RNA by using reverse transcription-PCR: sequence diversity within the lambda3-encoding L1 gene.";
RT J. Clin. Microbiol. 40:1368-1375(2002).
RL EMBL: AY007412; AAG28871.1; -.
DR HSSP; P17378; 1N35.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0019079; P:Viral genome replication; IEA.
FT NON_TER 1
FT NON_TER 121
SQ SEQUENCE 121 AA; 13391 MW; 4A53987DB0292B7C CRC64;

Query Match 64.4%; Score 38; DB 2; Length 121;
Best Local Similarity 75.0%; Pred. No. 28;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDD 8
|| |:|||

Db 96 NYVCGDD 103

RESULT 43

Q8V175 PRELIMINARY; PRT; 121 AA.
AC Q8V175;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Lambda 3 protein (fragment).
OS Reovirus sp.
OC Viruses; dsRNA viruses; Reoviridae; unclassified Reoviridae.
OX NCBI_TaxID=10891;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21920879; PubMed=11923358;
RA Leary T.P., Erker J.C., Chalmers M.L., Cruz A.T., Wetzel J.D.,
RA Desai S.M., Mushahwar I.K., Dermody T.S.;
RT "Detection of mammalian reovirus RNA by using reverse transcription-PCR: sequence diversity within the lambda3-encoding L1 gene.";
RT J. Clin. Microbiol. 40:1368-1375(2002).
RL EMBL: AY007411; AAG28870.1; -.
DR HSSP; P17378; 1N35.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0019079; P:Viral genome replication; IEA.
FT NON_TER 1
FT NON_TER 121
SQ SEQUENCE 121 AA; 13402 MW; 4A5D98791A892B7C CRC64;

Query Match 64.4%; Score 38; DB 2; Length 121;
Best Local Similarity 75.0%; Pred. No. 28;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Matches	6;	Conservative	1;	Mismatches	1;	Indels	0;	Gaps	0;
Qy	1	NYTCRGDD	8						
Db	96	NYVCQGDD	103						
RESULT 46									
Q8V178									
ID	Q8V178	PRELIMINARY;	PRT;	121	AA.				
AC	Q8V178;								
DT	01-MAR-2002	(TrEMBLrel. 20, Created)							
DT	01-MAR-2002	(TrEMBLrel. 20, Last sequence update)							
DT	01-JUN-2003	(TrEMBLrel. 24, Last annotation update)							
DE	Lambda 3 protein (Fragment).								
OS	Reovirus sp.								
OC	Viruses; dsRNA viruses; Reoviridae; unclassified Reoviridae.								
OX	NCBI_TaxID=10891;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RX	MEDLINE=21920879; PubMed=119233358;								
RA	Leary T.P., Erker J.C., Chalmers M.L., Cruz A.T., Wetzel J.D.,								
RS	Desai S.M., Muehahwar I.K., Dermody T.S.;								
RT	"Detection of mammalian reovirus RNA by using reverse transcription-								
RL	PCR: sequence diversity within the lambda3-encoding L1 gene.";								
RL	J. Clin. Microbiol. 40:1368-1375(2002).								
DR	EMBL; AY007408; BAG28867.1; -								
DR	HSP; P17378; IN35.								
DR	GO; GO:0003723; F:RNA binding; IEA.								
DR	GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.								
DR	GO; GO:0019079; P:Viral genome replication; IEA.								
FT	NON_TER	1							
FT	NON_TER	1							
FT	NON_TER	121							
SQ	SEQUENCE	121	AA;	13353	MM;	B947AB7396CDBB58	CRC64;		
Query Match 64.4%; Score 38; DB 2; Length 121;									
Best Local Similarity 75.0%; Pred. No. 28;									
Matches	6;	Conservative	1;	Mismatches	1;	Indels	0;	Gaps	0;

RESULT	47
ID	Q8V179
AC	PRELIMINARY; PRT; 121 AA.
DT	01-MAR-2002 (TrEMBLrel. 20, Created)
DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE	Lambda 3 protein (Fragment).
OS	Reovirus sp.
OC	Viruses; dsRNA viruses; Reoviridae; unclassified Reoviridae.
OX	NCBI_TaxID=10891;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=21920879; PubMed=11923358;
RA	Leary T.P., Erker J.C., Chalmers M.L., Cruz A.T., Wetzel J.D.,
RA	Desai S.M., Mushahwar I.K., Dermody T.S.;
RT	"Detection of mammalian reovirus RNA by using reverse transcription-
RT	PCR: sequence diversity within the lambda3-encoding l1 gene.";
RL	J. Clin. Microbiol. 40:1368-1375(2002).
DR	EMBL; AY007407; AAC2886.1; -.
DR	HSSP; P17378; 1N35.
DR	GO; GO:0003723; F:RNA binding; IEA.
DR	GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR	GO; GO:0019079; P:viral genome replication; IEA.
FT	NON_TER 1
FT	NON_TER 121
SQ	SEQUENCE 121 AA; 13377 MW; EA4498793D246CA CRC64;
Query Match	64.4%; Score 38; DB 2; Length 121;

Best Local Similarity 75.0%; Pred. No. 28;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 NYRCRGDD 8
DB 96 NYVCQGGD 103

RESULT 48

Q8V180 PRELIMINARY; PRT; 121 AA.
AC Q8V180;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Lambda 3 protein (Fragment).
OS Reovirus sp.
OC Viruses; dsRNA viruses; Reoviridae; unclassified Reoviridae.
OX NCBI_TaxID=10891;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21920879; PubMed=11923358;
RA Leary T.P., Erker J.C., Chalmers M.L., Cruz A.T., Wetzel J.D.,
RA Desai S.M., Mushahwar I.K., Dermody T.S.;
RT "Detection of mammalian reovirus RNA by using reverse transcription-
RT PCR: sequence diversity within the lambda3-encoding L1 gene.";
RL J. Clin. Microbiol. 40:1368-1375(2002).
DR EMBL; AY007406; BAG28865.1; -.
DR HSSP; P17378; IN35.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
FT NON_TER 1
FT NON_TER 121
SQ SEQUENCE 121 AA; 13377 MW; EA44987393D246CA CRC64;

Query Match 64.4%; Score 38; DB 2; Length 121;
Best Local Similarity 75.0%; Pred. No. 28;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDD 8
DB 96 NYVCQGGD 103

RESULT 49

Q8V181 PRELIMINARY; PRT; 121 AA.
AC Q8V181;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Lambda 3 protein (Fragment).
OS Reovirus sp.
OC Viruses; dsRNA viruses; Reoviridae; unclassified Reoviridae.
OX NCBI_TaxID=10891;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21920879; PubMed=11923358;
RA Leary T.P., Erker J.C., Chalmers M.L., Cruz A.T., Wetzel J.D.,
RA Desai S.M., Mushahwar I.K., Dermody T.S.;
RT "Detection of mammalian reovirus RNA by using reverse transcription-
RT PCR: sequence diversity within the lambda3-encoding L1 gene.";
RL J. Clin. Microbiol. 40:1368-1375(2002).
DR EMBL; AY007405; BAG28864.1; -.
DR HSSP; P17378; IN35.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
FT NON_TER 1
FT NON_TER 121
SQ SEQUENCE 121 AA; 13377 MW; EA44987393D246CA CRC64;

Query Match 64.4%; Score 38; DB 2; Length 121;
Best Local Similarity 75.0%; Pred. No. 28;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 NYRCRGDD 8
DB 96 NYVCQGGD 103

RESULT 50

Q8V182 PRELIMINARY; PRT; 121 AA.
AC Q8V182;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Lambda 3 protein (Fragment).
OS Reovirus sp.
OC Viruses; dsRNA viruses; Reoviridae; unclassified Reoviridae.
OX NCBI_TaxID=10891;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21920879; PubMed=11923358;
RA Leary T.P., Erker J.C., Chalmers M.L., Cruz A.T., Wetzel J.D.,
RA Desai S.M., Mushahwar I.K., Dermody T.S.;
RT "Detection of mammalian reovirus RNA by using reverse transcription-
RT PCR: sequence diversity within the lambda3-encoding L1 gene.";
RL J. Clin. Microbiol. 40:1368-1375(2002).
DR EMBL; AY007404; BAG28863.1; -.
DR HSSP; P17378; IN35.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
FT NON_TER 1
FT NON_TER 121
SQ SEQUENCE 121 AA; 13377 MW; EA44987393D246CA CRC64;

Query Match 64.4%; Score 38; DB 2; Length 121;
Best Local Similarity 75.0%; Pred. No. 28;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDD 8
DB 96 NYVCQGGD 103

Search completed: September 7, 2005, 19:57:27
Job time : 50.1163 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2005, 19:44:24 ; Search time 12.5581 Seconds
(without alignments)
59.443 Million cell updates/sec

Title: US-10-812-238B-20

Perfect score: 59

Sequence: 1 NYRCRDDSK 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/iaa/5A_COMB.pep:*
- 2: /cgn2_6/prodata/1/iaa/5B_COMB.pep:*
- 3: /cgn2_6/prodata/1/iaa/6A_COMB.pep:*
- 4: /cgn2_6/prodata/1/iaa/6B_COMB.pep:*
- 5: /cgn2_6/prodata/1/iaa/PCTUS_COMB.pep:*
- 6: /cgn2_6/prodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59	100.0	233	4	US-09-360-376-55
2	40	67.8	424	4	US-09-489-039A-12030
3	40	67.8	705	4	US-08-311-731A-4
4	39	66.1	501	4	US-09-270-767-41177
5	39	66.1	501	4	US-09-270-767-56393
6	37	62.7	275	4	US-09-489-039A-11789
7	37	62.7	488	4	US-09-489-039A-13164
8	36	61.0	416	4	US-09-949-016-8237
9	26	61.0	498	4	US-09-786-240-1
10	35	59.3	94	4	US-09-198-452A-1290
11	35	59.3	269	4	US-10-000-489-78
12	35	59.3	269	4	US-09-949-016-6121
13	35	59.3	276	4	US-09-949-016-7261
14	35	59.3	439	3	US-08-985-950-12
15	35	59.3	439	3	US-09-310-463-8
16	35	59.3	439	4	US-08-842-248A-8
17	35	59.3	439	4	US-09-546-049-12
18	35	59.3	503	4	US-09-248-796A-17110
19	35	59.3	621	2	US-08-419-652-4
20	35	59.3	660	1	US-08-248-532-3
21	35	59.3	660	2	US-08-419-652-3
22	35	59.3	662	1	US-08-248-532-2
23	35	59.3	662	2	US-08-419-652-2
24	35	59.3	662	2	US-08-685-118-4
25	35	59.3	662	2	US-08-915-495-4
26	35	59.3	662	2	US-08-914-520-4
27	35	59.3	662	3	US-08-789-350-2

28	59.3	662	4	US-09-949-016-6624	Sequence 5624, Ap
29	59.3	683	4	US-09-949-016-10092	Sequence 10092, A
30	59.3	700	4	US-09-252-991A-23746	Sequence 23746, A
31	57.6	163	4	US-09-252-991A-18617	Sequence 18617, A
32	57.6	272	4	US-09-489-039A-13748	Sequence 13748, A
33	57.6	280	4	US-09-601-478-5	Sequence 5, Appl
34	57.6	280	4	US-09-601-478-8	Sequence 8, Appl
35	57.6	280	4	US-09-442-013-13	Sequence 13, Appl
36	57.6	280	4	US-09-513-365A-1	Sequence 1, Appl
37	57.6	280	4	US-09-513-365A-6	Sequence 6, Appl
38	57.6	289	4	US-09-902-540-15729	Sequence 15729, A
39	57.6	316	4	US-09-583-110-4948	Sequence 4948, Ap
40	57.6	321	4	US-09-107-433-5053	Sequence 5053, Ap
41	57.6	328	4	US-09-252-991A-30671	Sequence 30671, A
42	57.6	357	4	US-09-248-796A-21669	Sequence 21669, A
43	57.6	447	3	US-09-627-376-10	Sequence 10, Appl
44	57.6	447	4	US-10-047-676B-10	Sequence 10, Appl
45	57.6	971	3	US-09-107-149-19	Sequence 19, Appl
46	57.6	1036	3	US-08-968-752B-4	Sequence 4, Appl
47	57.6	1036	3	US-09-536-224-4	Sequence 4, Appl
48	57.6	1074	2	US-08-768-147B-2	Sequence 2, Appl
49	57.6	1074	3	US-08-968-752B-2	Sequence 2, Appl
50	57.6	1074	3	US-09-107-149-3	Sequence 3, Appl
51	57.6	1074	3	US-09-536-224-2	Sequence 2, Appl
52	57.6	1086	4	US-09-949-016-11159	Sequence 11159, A
53	55.9	103	4	US-09-107-532A-7072	Sequence 7072, Ap
54	55.9	158	4	US-09-248-796A-18178	Sequence 18178, A
55	55.9	174	4	US-09-107-532A-4715	Sequence 4715, Ap
56	55.9	201	3	US-09-015-734-12	Sequence 12, Appl
57	55.9	201	4	US-09-515-311-12	Sequence 12, Appl
58	55.9	236	3	US-09-015-734-7	Sequence 7, Appl
59	55.9	236	4	US-09-515-311-7	Sequence 7, Appl
60	55.9	253	3	US-09-252-991A-32459	Sequence 32459, A
61	55.9	255	3	US-09-015-734-2	Sequence 2, Appl
62	55.9	255	4	US-09-515-311-2	Sequence 2, Appl
63	55.9	272	1	US-08-282-951-6	Sequence 6, Appl
64	55.9	299	4	US-09-252-991A-21857	Sequence 21857, A
65	55.9	389	4	US-09-270-767-32291	Sequence 32291, A
66	55.9	389	4	US-09-270-767-47508	Sequence 47508, A
67	55.9	503	4	US-09-902-540-11343	Sequence 11343, A
68	55.9	686	4	US-09-252-991A-18115	Sequence 18115, A
69	55.9	730	1	US-08-121-713D-58	Sequence 58, Appl
70	55.9	730	1	US-08-835-268-58	Sequence 58, Appl
71	55.9	730	2	US-09-060-692-58	Sequence 58, Appl
72	55.9	730	3	US-08-833-391-58	Sequence 58, Appl
73	55.9	730	3	US-09-060-610-58	Sequence 58, Appl
74	55.9	730	5	PCT-US94-10151A-58	Sequence 58, Appl
75	55.9	800	4	US-09-489-039A-7191	Sequence 7191, Ap
76	55.9	819	4	US-09-949-016-11044	Sequence 11044, A
77	55.9	825	4	US-09-489-039A-8469	Sequence 8469, Ap
78	55.9	837	4	US-09-949-016-6515	Sequence 6515, Ap
79	54.2	7	1	US-08-421-702A-17	Sequence 17, Appl
80	54.2	7	1	US-08-303-052A-17	Sequence 17, Appl
81	54.2	7	1	US-08-421-696A-17	Sequence 17, Appl
82	54.2	7	1	US-08-421-697A-17	Sequence 17, Appl
83	54.2	7	1	US-08-421-698A-17	Sequence 17, Appl
84	54.2	7	5	PCT-US95-04741-17	Sequence 17, Appl
85	54.2	9	1	US-08-421-702A-26	Sequence 26, Appl
86	54.2	9	1	US-08-421-696A-26	Sequence 26, Appl
87	54.2	9	1	US-08-421-697A-26	Sequence 26, Appl
88	54.2	9	1	US-08-421-698A-26	Sequence 26, Appl
89	54.2	9	1	US-08-421-699A-26	Sequence 26, Appl
90	54.2	9	5	PCT-US95-04741-26	Sequence 26, Appl
91	54.2	20	3	US-08-857-076-27	Sequence 27, Appl
92	54.2	35	3	US-08-857-076-30	Sequence 30, Appl
93	54.2	43	3	US-08-857-076-28	Sequence 28, Appl
94	54.2	62	4	US-09-248-796A-35139	Sequence 25139, A
95	54.2	90	4	US-09-270-767-31781	Sequence 31781, A
96	54.2	90	4	US-09-270-767-38050	Sequence 38050, A
97	54.2	90	4	US-09-270-767-46998	Sequence 46998, A
98	54.2	90	4	US-09-270-767-53267	Sequence 53267, A
99	54.2	98	4	US-09-513-999C-4239	Sequence 4239, Ap
100	54.2	98	4		

101	32	54.2	107	4	US-09-107-532A-5073	Sequence 5073, Ap	174	32	54.2	171	5	PCT-US95-00454-24	Sequence 24, Appl
102	32	54.2	118	4	US-08-794-028B-21	Sequence 21, Appl	175	32	54.2	214	3	US-08-846-790A-1	Sequence 1, Appl
103	32	54.2	118	4	US-08-794-028B-22	Sequence 22, Appl	176	32	54.2	214	3	US-08-935-333-1	Sequence 1, Appl
104	32	54.2	119	3	US-08-753-642-2	Sequence 2, Appl	177	32	54.2	229	4	US-09-252-991A-22195	Sequence 22195, A
105	32	54.2	120	1	US-07-979-630-1	Sequence 1, Appl	178	32	54.2	241	1	US-08-266-080B-4	Sequence 4, Appl
106	32	54.2	120	1	US-08-440-049-1	Sequence 1, Appl	179	32	54.2	241	1	US-08-451-947-5	Sequence 5, Appl
107	32	54.2	120	1	US-08-440-049-3	Sequence 3, Appl	180	32	54.2	241	2	US-08-424-826A-5	Sequence 5, Appl
108	32	54.2	120	1	US-08-441-513A-1	Sequence 1, Appl	181	32	54.2	241	2	US-08-595-043A-75	Sequence 75, Appl
109	32	54.2	120	2	US-08-441-513A-3	Sequence 3, Appl	182	32	54.2	241	3	US-08-970-865-1	Sequence 1, Appl
110	32	54.2	120	3	US-08-970-865-2	Sequence 2, Appl	183	32	54.2	241	3	US-08-928-694-5	Sequence 5, Appl
111	32	54.2	120	3	US-08-970-865-3	Sequence 3, Appl	184	32	54.2	241	3	US-09-363-573-1	Sequence 1, Appl
112	32	54.2	120	3	US-08-581-662-31	Sequence 31, Appl	185	32	54.2	241	3	US-09-447-356-3	Sequence 3, Appl
113	32	54.2	120	3	US-09-363-573-2	Sequence 2, Appl	186	32	54.2	241	4	US-08-450-842-5	Sequence 5, Appl
114	32	54.2	120	3	US-09-363-573-3	Sequence 3, Appl	187	32	54.2	241	4	US-09-620-174-2	Sequence 2, Appl
115	32	54.2	120	3	US-08-845-541B-1	Sequence 1, Appl	188	32	54.2	241	4	US-09-620-174-2	Sequence 2, Appl
116	32	54.2	120	3	US-08-845-541B-3	Sequence 3, Appl	189	32	54.2	241	5	PCT-US91-06950-5	Sequence 5, Appl
117	32	54.2	120	3	US-08-845-541B-4	Sequence 4, Appl	190	32	54.2	241	5	PCT-US95-05423-4	Sequence 4, Appl
118	32	54.2	120	3	US-08-845-541B-5	Sequence 5, Appl	191	32	54.2	242	4	US-09-675-503-1	Sequence 1, Appl
119	32	54.2	120	3	US-08-845-541B-6	Sequence 6, Appl	192	32	54.2	242	4	US-10-072-681-1	Sequence 1, Appl
120	32	54.2	120	3	US-08-845-541B-7	Sequence 7, Appl	193	32	54.2	247	4	US-09-368-819A-2	Sequence 2, Appl
121	32	54.2	120	3	US-08-845-541B-8	Sequence 8, Appl	194	32	54.2	264	4	US-09-368-819A-4	Sequence 4, Appl
122	32	54.2	120	3	US-08-845-541B-9	Sequence 9, Appl	195	32	54.2	280	4	US-09-252-991A-18217	Sequence 18217, A
123	32	54.2	120	3	US-08-845-541B-10	Sequence 10, Appl	196	32	54.2	283	4	US-09-252-991A-32946	Sequence 32946, A
124	32	54.2	120	3	US-08-845-541B-11	Sequence 11, Appl	197	32	54.2	316	4	US-09-328-352-8163	Sequence 8163, Ap
125	32	54.2	120	3	US-08-845-541B-12	Sequence 12, Appl	198	32	54.2	350	4	US-09-675-305-12	Sequence 12, Appl
126	32	54.2	120	3	US-08-845-541B-13	Sequence 13, Appl	199	32	54.2	350	4	US-10-200-344-12	Sequence 12, Appl
127	32	54.2	120	3	US-08-845-541B-17	Sequence 17, Appl	200	32	54.2	369	4	US-09-252-991A-25394	Sequence 25394, A
128	32	54.2	120	3	US-08-845-541B-18	Sequence 18, Appl	201	32	54.2	376	4	US-09-489-039A-8789	Sequence 8789, Ap
129	32	54.2	120	3	US-08-845-541B-19	Sequence 19, Appl	202	32	54.2	390	4	US-08-993-380-4	Sequence 4, Appl
130	32	54.2	120	3	US-08-845-541B-20	Sequence 20, Appl	203	32	54.2	410	4	US-09-252-991A-31174	Sequence 31174, A
131	32	54.2	120	3	US-08-845-541B-21	Sequence 21, Appl	204	32	54.2	437	4	US-09-675-305-10	Sequence 10, Appl
132	32	54.2	120	3	US-09-066-065A-1	Sequence 1, Appl	205	32	54.2	437	4	US-10-200-344-10	Sequence 10, Appl
133	32	54.2	120	3	US-09-066-065A-3	Sequence 3, Appl	206	32	54.2	508	4	US-09-858-664A-18	Sequence 18, Appl
134	32	54.2	120	3	US-09-066-065A-4	Sequence 4, Appl	207	32	54.2	508	4	US-10-274-978-19	Sequence 19, Appl
135	32	54.2	120	3	US-09-066-065A-5	Sequence 5, Appl	208	32	54.2	508	4	US-10-697-263-19	Sequence 19, Appl
136	32	54.2	120	3	US-09-066-065A-6	Sequence 6, Appl	209	32	54.2	539	2	US-08-978-182-3	Sequence 3, Appl
137	32	54.2	120	3	US-09-066-065A-7	Sequence 7, Appl	210	32	54.2	539	2	US-09-205-681-3	Sequence 3, Appl
138	32	54.2	120	3	US-09-066-065A-8	Sequence 8, Appl	211	32	54.2	556	4	US-09-252-991A-26282	Sequence 26282, A
139	32	54.2	120	3	US-09-066-065A-9	Sequence 9, Appl	212	32	54.2	714	4	US-09-328-352-5347	Sequence 5347, Ap
140	32	54.2	120	3	US-09-066-065A-10	Sequence 10, Appl	213	32	54.2	1429	1	US-07-642-002-2	Sequence 2, Appl
141	32	54.2	120	3	US-09-066-065A-11	Sequence 11, Appl	214	32	54.2	1429	2	US-08-365-486A-13	Sequence 13, Appl
142	32	54.2	120	3	US-09-066-065A-12	Sequence 12, Appl	215	32	54.2	1429	2	US-08-319-866-11	Sequence 11, Appl
143	32	54.2	120	3	US-09-066-065A-13	Sequence 13, Appl	216	32	54.2	1429	3	US-08-880-342-13	Sequence 13, Appl
144	32	54.2	120	3	US-09-066-065A-17	Sequence 17, Appl	217	32	54.2	1429	4	US-09-661-258-2	Sequence 2, Appl
145	32	54.2	120	3	US-09-066-065A-18	Sequence 18, Appl	218	32	54.2	1429	4	US-08-809-917-11	Sequence 11, Appl
146	32	54.2	120	3	US-09-066-065A-19	Sequence 19, Appl	219	32	54.2	1430	2	US-08-705-625-4	Sequence 4, Appl
147	32	54.2	120	3	US-09-066-065A-20	Sequence 20, Appl	220	32	54.2	1430	3	US-09-010-998-5	Sequence 5, Appl
148	32	54.2	120	3	US-09-066-065A-21	Sequence 21, Appl	221	32	54.2	1430	3	US-09-220-574-4	Sequence 4, Appl
149	32	54.2	120	3	US-09-447-356-1	Sequence 1, Appl	222	32	54.2	1433	2	US-08-365-486A-21	Sequence 21, Appl
150	32	54.2	120	4	US-09-664-295-31	Sequence 31, Appl	223	32	54.2	1433	3	US-09-123-708-4	Sequence 4, Appl
151	32	54.2	120	4	US-08-794-028B-1	Sequence 1, Appl	224	32	54.2	1433	3	US-09-123-624-4	Sequence 4, Appl
152	32	54.2	120	4	US-08-794-028B-3	Sequence 3, Appl	225	32	54.2	1433	3	US-08-880-342-21	Sequence 21, Appl
153	32	54.2	120	5	PCT-US93-11282-1	Sequence 1, Appl	226	32	54.2	1434	2	US-08-365-486A-19	Sequence 19, Appl
154	32	54.2	120	5	PCT-US95-06918-1	Sequence 1, Appl	227	32	54.2	1434	3	US-08-880-342-19	Sequence 19, Appl
155	32	54.2	120	5	PCT-US95-06918-3	Sequence 3, Appl	228	32	54.2	1434	4	US-09-661-258-1	Sequence 1, Appl
156	32	54.2	121	4	US-09-675-503-2	Sequence 2, Appl	229	32	54.2	1434	4	US-09-538-092-1034	Sequence 1034, Ap
157	32	54.2	121	4	US-09-675-503-3	Sequence 3, Appl	230	32	54.2	1529	3	US-09-215-694-1	Sequence 1, Appl
158	32	54.2	121	4	US-10-072-681-2	Sequence 2, Appl	231	32	54.2	1554	3	US-08-705-625-3	Sequence 3, Appl
159	32	54.2	121	4	US-10-072-681-3	Sequence 3, Appl	232	32	54.2	1554	3	US-09-010-998-6	Sequence 6, Appl
160	32	54.2	125	4	US-09-489-039A-11259	Sequence 11259, A	233	32	54.2	1554	3	US-09-220-574-3	Sequence 3, Appl
161	32	54.2	153	4	US-09-675-922-2	Sequence 2, Appl	234	32	54.2	9	1	US-08-482-106-12	Sequence 12, Appl
162	32	54.2	155	4	US-09-270-767-36322	Sequence 36322, A	235	31	52.5	45	6	5240706-3	Patent No. 5240706
163	32	54.2	155	4	US-09-270-767-51539	Sequence 51539, A	236	31	52.5	45	6	5240706-3	Patent No. 5240706
164	32	54.2	157	4	US-09-675-922-4	Sequence 4, Appl	237	31	52.5	98	4	US-09-198-452A-1291	Sequence 1291, Ap
165	32	54.2	163	4	US-09-675-922-6	Sequence 6, Appl	238	31	52.5	100	4	US-09-621-976-6277	Sequence 6277, Ap
166	32	54.2	164	4	US-09-949-016-10989	Sequence 10989, A	239	31	52.5	107	4	US-09-484-577A-86	Sequence 86, Appl
167	32	54.2	167	4	US-09-675-922-8	Sequence 8, Appl	240	31	52.5	115	4	US-09-270-767-40288	Sequence 40288, A
168	32	54.2	171	2	US-08-417-495-24	Sequence 24, Appl	241	31	52.5	115	4	US-09-270-767-55504	Sequence 55504, A
169	32	54.2	171	2	US-08-284-391B-24	Sequence 24, Appl	242	31	52.5	135	4	US-09-489-039A-9712	Sequence 9712, Ap
170	32	54.2	171	3	US-09-218-950-24	Sequence 24, Appl	243	31	52.5	154	4	US-09-270-767-46510	Sequence 46510, A
171	32	54.2	171	4	US-08-394-388A-24	Sequence 24, Appl	244	31	52.5	182	2	US-08-417-495-25	Sequence 25, Appl
172	32	54.2	171	4	US-09-949-016-16669	Sequence 16669, A	245	31	52.5	182	2	US-08-284-391B-25	Sequence 25, Appl
173	32	54.2	171	5	PCT-US92-01785-24	Sequence 24, Appl	246	31	52.5	182	3	US-09-218-950-25	Sequence 25, Appl

247	31	52.5	182	4	US-08-394-388A-25	Sequence 25, Appl	320	30	50.8	160	2	US-08-448-438-6	Sequence 6, Appli
248	31	52.5	182	5	PCT-US92-01785-25	Sequence 25, Appl	321	30	50.8	160	2	US-08-448-438-7	Sequence 7, Appli
249	31	52.5	182	5	PCT-US95-00454-25	Sequence 25, Appl	322	30	50.8	167	4	US-09-270-767-45525	Sequence 45525, A
250	31	52.5	184	4	US-09-252-991A-17582	Sequence 17582, A	323	30	50.8	172	4	US-09-252-991A-22192	Sequence 22192, A
251	31	52.5	194	4	US-09-949-016-8204	Sequence 8204, Ap	324	30	50.8	190	4	US-09-248-796A-16487	Sequence 16487, A
252	31	52.5	235	4	US-09-533-029-74	Sequence 74, Appl	325	30	50.8	202	4	US-09-248-796A-27997	Sequence 27997, A
253	31	52.5	249	4	US-09-302-540-13089	Sequence 13089, A	326	30	50.8	206	4	US-09-270-767-57339	Sequence 57339, A
254	31	52.5	254	4	US-09-266-965-123	Sequence 123, App	327	30	50.8	210	2	US-09-258-257-2	Sequence 2, Appli
255	31	52.5	265	4	US-09-252-991A-32218	Sequence 32218, A	328	30	50.8	210	2	US-09-258-371-2	Sequence 2, Appli
256	31	52.5	301	4	US-09-252-991A-17405	Sequence 17405, A	329	30	50.8	210	3	US-08-569-721A-2	Sequence 2, Appli
257	31	52.5	302	4	US-08-311-731A-206	Sequence 206, App	330	30	50.8	210	3	US-08-751-230-2	Sequence 2, Appli
258	31	52.5	338	4	US-09-248-796A-17967	Sequence 17967, A	331	30	50.8	210	3	US-09-499-082-2	Sequence 2, Appli
259	31	52.5	340	4	US-08-489-039A-11464	Sequence 11464, A	332	30	50.8	210	3	US-09-258-372-2	Sequence 2, Appli
260	31	52.5	341	4	US-09-252-991A-16890	Sequence 16890, A	333	30	50.8	210	3	US-09-006-783A-7	Sequence 7, Appli
261	31	52.5	347	3	US-09-259-369-2	Sequence 2, Appli	334	30	50.8	210	3	US-09-159-871-13	Sequence 13, Appl
262	31	52.5	364	4	US-09-328-352-4659	Sequence 4659, Ap	335	30	50.8	210	4	US-09-532-868-2	Sequence 2, Appli
263	31	52.5	374	4	US-09-252-991A-17863	Sequence 17863, A	336	30	50.8	210	4	US-09-451-739H-6	Sequence 6, Appli
264	31	52.5	374	4	US-09-252-991A-23565	Sequence 23565, A	337	30	50.8	217	4	US-09-369-247-113	Sequence 113, App
265	31	52.5	385	4	US-09-489-039A-10159	Sequence 10159, A	338	30	50.8	225	4	US-09-673-395A-354	Sequence 354, App
266	31	52.5	401	3	US-09-005-298-17	Sequence 17, Appl	339	30	50.8	230	4	US-09-252-991A-22765	Sequence 7, Appli
267	31	52.5	401	5	PCT-US96-09848-31	Sequence 31, Appl	340	30	50.8	235	4	US-09-451-739H-7	Sequence 7, Appli
268	31	52.5	401	5	US-08-768-619-17	Sequence 17, Appl	341	30	50.8	250	4	US-09-270-767-43235	Sequence 43235, A
269	31	52.5	407	3	US-09-005-298-15	Sequence 15, Appl	342	30	50.8	268	4	US-09-489-039A-13115	Sequence 13115, A
270	31	52.5	407	4	US-08-768-619-15	Sequence 15, Appl	343	30	50.8	269	3	US-09-006-783A-5	Sequence 5, Appli
271	31	52.5	407	5	PCT-US96-09848-15	Sequence 15, Appl	344	30	50.8	269	4	US-09-159-871-11	Sequence 11, Appl
272	31	52.5	450	4	US-09-252-991A-18376	Sequence 18376, A	345	30	50.8	270	3	US-09-085-305-16	Sequence 16, Appl
273	31	52.5	450	4	US-09-252-991A-25835	Sequence 25835, A	346	30	50.8	270	3	US-09-085-305-19	Sequence 19, Appl
274	31	52.5	475	4	US-09-302-540-15337	Sequence 15337, A	347	30	50.8	278	4	US-09-248-796A-16036	Sequence 16036, A
275	31	52.5	495	2	US-08-794-795-2	Sequence 2, Appli	348	30	50.8	279	4	US-09-252-991A-19165	Sequence 19165, A
276	31	52.5	495	3	US-09-249-200-2	Sequence 2, Appli	349	30	50.8	279	4	US-09-442-013-11	Sequence 11, Appl
277	31	52.5	520	2	US-08-794-795-6	Sequence 6, Appli	350	30	50.8	279	4	US-09-451-739H-5	Sequence 5, Appli
278	31	52.5	520	3	US-09-249-200-6	Sequence 6, Appli	351	30	50.8	279	4	US-09-513-365A-8	Sequence 8, Appli
279	31	52.5	521	3	US-09-252-991A-18266	Sequence 18266, A	352	30	50.8	282	4	US-09-360-376-54	Sequence 9, Appli
280	31	52.5	562	3	US-09-012-515A-14	Sequence 14, Appl	353	30	50.8	282	4	US-09-360-376-54	Sequence 54, Appl
281	31	52.5	562	3	US-08-360-144A-14	Sequence 14, Appl	354	30	50.8	283	4	US-09-270-767-41831	Sequence 41831, A
282	31	52.5	562	4	US-09-012-504A-14	Sequence 14, Appl	355	30	50.8	294	2	US-09-258-371-10	Sequence 10, Appl
283	31	52.5	562	4	US-09-012-399A-14	Sequence 14, Appl	356	30	50.8	294	3	US-08-751-230-10	Sequence 10, Appl
284	31	52.5	562	5	PCT-US95-06722-14	Sequence 14, Appl	357	30	50.8	294	3	US-09-499-082-10	Sequence 10, Appl
285	31	52.5	615	4	US-09-538-092-615	Sequence 615, App	358	30	50.8	294	3	US-09-258-372-10	Sequence 10, Appl
286	31	52.5	625	4	US-09-252-991A-19871	Sequence 19871, A	359	30	50.8	294	3	US-09-006-783A-3	Sequence 3, Appli
287	31	52.5	718	4	US-09-252-991A-25696	Sequence 25696, A	360	30	50.8	294	4	US-09-159-871-2	Sequence 2, Appli
288	30	50.8	9	1	US-08-482-106-11	Sequence 11, Appl	361	30	50.8	294	4	US-09-442-013-12	Sequence 12, Appl
289	30	50.8	24	4	US-09-369-247-118	Sequence 118, App	362	30	50.8	294	4	US-09-532-868-10	Sequence 10, Appl
290	30	50.8	29	2	US-08-620-151-71	Sequence 71, Appl	363	30	50.8	294	4	US-09-451-739H-19	Sequence 19, Appl
291	30	50.8	29	3	US-09-253-396A-212	Sequence 212, App	364	30	50.8	297	4	US-09-902-540-14353	Sequence 14353, A
292	30	50.8	39	1	US-07-602-847C-16	Sequence 16, Appl	365	30	50.8	301	4	US-09-902-540-11985	Sequence 11985, A
293	30	50.8	39	3	US-08-817-787-32	Sequence 32, Appl	366	30	50.8	308	4	US-09-252-991A-27921	Sequence 27921, A
294	30	50.8	39	4	US-09-583-808-32	Sequence 32, Appl	367	30	50.8	321	3	US-08-485-511A-4	Sequence 4, Appli
295	30	50.8	41	3	US-08-817-787-36	Sequence 36, Appl	368	30	50.8	357	4	US-09-949-016-8302	Sequence 8302, Ap
296	30	50.8	41	4	US-09-583-808-36	Sequence 36, Appl	369	30	50.8	360	2	US-09-252-991A-29934	Sequence 29934, A
297	30	50.8	42	3	US-08-817-787-34	Sequence 34, Appl	370	30	50.8	375	2	US-08-506-864A-1	Sequence 1, Appli
298	30	50.8	42	4	US-08-358-160-126	Sequence 126, App	371	30	50.8	377	4	US-08-851-968-1	Sequence 1, Appli
299	30	50.8	57	1	US-08-358-160-126	Sequence 126, App	372	30	50.8	377	4	US-09-949-016-8302	Sequence 8302, Ap
300	30	50.8	60	4	US-08-469-260A-450	Sequence 450, App	373	30	50.8	379	4	US-09-345-828-2	Sequence 2, Appli
301	30	50.8	60	4	US-08-488-446-450	Sequence 450, App	374	30	50.8	389	4	US-09-949-016-9951	Sequence 9951, Ap
302	30	50.8	60	4	US-08-467-344A-450	Sequence 450, App	375	30	50.8	395	4	US-09-949-016-9496	Sequence 9496, Ap
303	30	50.8	60	4	US-08-424-550B-450	Sequence 450, App	376	30	50.8	406	4	US-09-543-681A-7962	Sequence 7962, Ap
304	30	50.8	92	4	US-09-302-540-14307	Sequence 14307, A	377	30	50.8	412	4	US-09-252-991A-30352	Sequence 30352, A
305	30	50.8	97	4	US-09-370-767-57075	Sequence 57075, A	378	30	50.8	416	4	US-09-949-016-6648	Sequence 6648, Ap
306	30	50.8	99	4	US-09-621-976-4220	Sequence 4220, Ap	379	30	50.8	417	3	US-08-705-771-18	Sequence 18, Appl
307	30	50.8	132	4	US-09-252-991A-30452	Sequence 30452, A	380	30	50.8	417	4	US-09-417-540-18	Sequence 18, Appl
308	30	50.8	137	3	US-09-188-930-174	Sequence 174, App	381	30	50.8	428	1	US-07-973-324A-4	Sequence 4, Appli
309	30	50.8	137	4	US-09-312-283C-174	Sequence 174, App	382	30	50.8	428	1	US-08-343-380-4	Sequence 4, Appli
310	30	50.8	141	1	US-07-695-564-5	Sequence 5, Appli	383	30	50.8	428	3	US-09-072-435-4	Sequence 4, Appli
311	30	50.8	141	1	US-08-241-387-5	Sequence 5, Appli	384	30	50.8	428	3	US-09-072-917A-4	Sequence 4, Appli
312	30	50.8	142	2	US-08-448-438-1	Sequence 1, Appli	385	30	50.8	433	3	US-09-697-367-22	Sequence 22, Appl
313	30	50.8	142	2	US-08-448-438-2	Sequence 2, Appli	386	30	50.8	433	3	US-09-918-909A-22	Sequence 22, Appl
314	30	50.8	142	2	US-08-448-438-3	Sequence 3, Appli	387	30	50.8	434	1	US-07-923-692C-6	Sequence 6, Appli
315	30	50.8	149	1	US-07-695-564-7	Sequence 7, Appli	388	30	50.8	434	2	US-08-184-237-6	Sequence 6, Appli
316	30	50.8	149	1	US-08-241-387-7	Sequence 7, Appli	389	30	50.8	434	2	US-08-482-920-6	Sequence 6, Appli
317	30	50.8	156	1	US-07-766-682A-3	Sequence 3, Appli	390	30	50.8	434	3	US-08-484-341-6	Sequence 6, Appli
318	30	50.8	156	4	US-09-398-412B-6	Sequence 6, Appli	391	30	50.8	434	3	US-08-483-502-6	Sequence 6, Appli
319	30	50.8	160	2	US-08-448-438-5	Sequence 5, Appli	392	30	50.8	434	4	US-09-726-651A-6	Sequence 6, Appli

393	30	50.8	437	4	US-09-918-909A-28	Sequence 28, Appl	466	29	49.2	80	4	US-09-134-000C-5829	Sequence 5829, Ap
394	30	50.8	440	4	US-09-302-540-9852	Sequence 9752, Ap	467	29	49.2	82	4	US-09-134-000C-6553	Sequence 6553, Ap
395	30	50.8	441	4	US-09-949-016-9778	Sequence 9878, Ap	468	29	49.2	83	4	US-09-107-433-4178	Sequence 4178, Ap
396	30	50.8	461	4	US-09-252-991A-28623	Sequence 28623, A	469	29	49.2	99	4	US-09-489-039A-7470	Sequence 7470, Ap
397	30	50.8	469	4	US-09-452-991A-26584	Sequence 26584, A	470	29	49.2	101	1	US-08-335-583C-54	Sequence 54, Appl
398	30	50.8	471	4	US-09-302-540-11973	Sequence 11973, A	471	29	49.2	101	3	US-08-931-858E-155	Sequence 155, App
399	30	50.8	483	4	US-09-305-999-20	Sequence 20, Appl	472	29	49.2	101	3	US-08-981-739-155	Sequence 155, App
400	30	50.8	502	2	US-08-459-818-19	Sequence 19, Appl	473	29	49.2	101	4	US-09-128-026-155	Sequence 155, App
401	30	50.8	502	2	US-08-889-666-19	Sequence 19, Appl	474	29	49.2	101	4	US-09-220-616-155	Sequence 155, App
402	30	50.8	502	2	US-08-465-078-19	Sequence 19, Appl	475	29	49.2	101	4	US-09-220-527-155	Sequence 155, App
403	30	50.8	502	2	US-08-725-776-19	Sequence 19, Appl	476	29	49.2	101	4	US-09-220-407-155	Sequence 155, App
404	30	50.8	502	2	US-08-488-062-19	Sequence 19, Appl	477	29	49.2	105	4	US-09-489-039A-13562	Sequence 13562, A
405	30	50.8	517	4	US-09-252-991A-22249	Sequence 22249, A	478	29	49.2	118	3	US-08-624-635-21	Sequence 21, Appl
406	30	50.8	596	4	US-09-302-540-13547	Sequence 13547, A	479	29	49.2	120	4	US-09-513-999C-6407	Sequence 6407, Ap
407	30	50.8	597	4	US-09-949-016-7800	Sequence 7800, Ap	480	29	49.2	122	4	US-09-775-932-10	Sequence 10, Appl
408	30	50.8	615	4	US-09-540-236-2639	Sequence 2639, Ap	481	29	49.2	124	4	US-09-252-991A-29290	Sequence 29290, A
409	30	50.8	673	4	US-09-538-092-1270	Sequence 1270, Ap	482	29	49.2	124	4	US-09-621-976-4699	Sequence 4699, Ap
410	30	50.8	717	6	5262177-5	Patent No. 5262177	483	29	49.2	126	2	US-08-482-728A-15	Sequence 15, Appl
411	30	50.8	717	6	5262177-5	Patent No. 5262177	484	29	49.2	126	4	US-08-849-303-27	Sequence 27, Appl
412	30	50.8	719	2	US-08-520-933-3	Sequence 3, Appli	485	29	49.2	131	4	US-09-621-976-5011	Sequence 5011, Ap
413	30	50.8	719	4	US-09-285-040-3	Sequence 3, Appli	486	29	49.2	136	4	US-09-690-454-131	Sequence 131, App
414	30	50.8	738	6	5262177-2	Patent No. 5262177	487	29	49.2	137	4	US-09-270-767-59045	Sequence 59045, A
415	30	50.8	738	6	5262177-2	Patent No. 5262177	488	29	49.2	138	4	US-09-540-236-2621	Sequence 2621, Ap
416	30	50.8	776	4	US-09-949-016-10829	Sequence 10829, A	489	29	49.2	138	4	US-09-270-767-34849	Sequence 34849, A
417	30	50.8	777	4	US-09-270-767-44409	Sequence 44409, A	490	29	49.2	138	4	US-09-270-767-50066	Sequence 50066, A
418	30	50.8	911	4	US-09-252-991A-25237	Sequence 25237, A	491	29	49.2	139	4	US-09-270-767-34282	Sequence 34282, A
419	30	50.8	943	4	US-09-623-624-4	Sequence 4, Appli	492	29	49.2	139	4	US-09-270-767-49499	Sequence 49499, A
420	30	50.8	943	4	US-10-270-595-4	Sequence 4, Appli	493	29	49.2	142	3	US-08-744-138-4	Sequence 4, Appli
421	30	50.8	964	4	US-09-556-877-177	Sequence 177, App	494	29	49.2	142	3	US-09-431-480-7	Sequence 7, Appli
422	30	50.8	964	4	US-09-620-413C-177	Sequence 177, App	495	29	49.2	142	3	US-09-617-302-7	Sequence 7, Appli
423	30	50.8	964	4	US-09-598-419-177	Sequence 177, App	496	29	49.2	142	3	US-09-241-376-4	Sequence 4, Appli
424	30	50.8	977	4	US-09-556-877-191	Sequence 191, App	497	29	49.2	142	4	US-09-940-497-4	Sequence 4, Appli
425	30	50.8	977	4	US-09-620-412C-191	Sequence 191, App	498	29	49.2	142	4	US-09-976-594-358	Sequence 358, App
426	30	50.8	977	4	US-09-598-419-191	Sequence 191, App	499	29	49.2	142	4	US-08-849-303-20	Sequence 20, Appl
427	30	50.8	1073	1	US-07-695-564-1	Sequence 1, Appli	500	29	49.2	145	2	US-08-788-584-5	Sequence 5, Appli
428	30	50.8	1073	1	US-08-241-387-1	Sequence 1, Appli							
429	30	50.8	1091	1	US-07-695-564-3	Sequence 3, Appli							
430	30	50.8	1091	1	US-08-241-387-3	Sequence 3, Appli							
431	30	50.8	1205	1	US-09-491-522-11	Sequence 11, Appl							
432	30	50.8	1248	4	US-09-949-016-10595	Sequence 10595, A							
433	30	50.8	1248	4	US-09-949-016-10596	Sequence 10596, A							
434	30	50.8	1338	4	US-09-949-016-6029	Sequence 6029, Ap							
435	30	50.8	3924	4	US-09-538-092-1246	Sequence 1246, Ap							
436	30	50.8	4544	1	US-08-469-486-52	Sequence 52, Appl							
437	30	50.8	4544	2	US-08-469-658-52	Sequence 52, Appl							
438	29	49.2	7	1	US-08-421-702A-8	Sequence 8, Appli							
439	29	49.2	7	1	US-08-421-696A-8	Sequence 8, Appli							
440	29	49.2	7	1	US-08-421-697A-8	Sequence 8, Appli							
441	29	49.2	7	1	US-08-421-698A-8	Sequence 8, Appli							
442	29	49.2	7	1	US-08-421-695A-8	Sequence 8, Appli							
443	29	49.2	7	5	PCT-US95-04741-8	Sequence 8, Appli							
444	29	49.2	34	4	US-09-270-767-34502	Sequence 34502, A							
445	29	49.2	34	4	US-09-270-767-49719	Sequence 49719, A							
446	29	49.2	46	3	US-09-257-218-16	Sequence 16, Appl							
447	29	49.2	46	3	US-09-311-760-16	Sequence 16, Appl							
448	29	49.2	46	4	US-08-865-579-16	Sequence 16, Appl							
449	29	49.2	46	4	US-10-059-749-16	Sequence 16, Appl							
450	29	49.2	56	3	US-08-711-417C-186	Sequence 186, App							
451	29	49.2	56	3	US-08-711-417C-187	Sequence 187, App							
452	29	49.2	56	3	US-08-711-417C-188	Sequence 188, App							
453	29	49.2	56	3	US-08-711-417C-189	Sequence 189, App							
454	29	49.2	56	3	US-09-187-789-58	Sequence 58, Appl							
455	29	49.2	56	4	US-09-139-600-53	Sequence 53, Appl							
456	29	49.2	56	4	US-09-723-909-186	Sequence 186, App							
457	29	49.2	56	4	US-09-723-909-187	Sequence 187, App							
458	29	49.2	56	4	US-09-723-909-188	Sequence 188, App							
459	29	49.2	56	4	US-09-723-909-189	Sequence 189, App							
460	29	49.2	56	4	US-09-989-903-58	Sequence 58, Appl							
461	29	49.2	64	4	US-09-328-352-5321	Sequence 5321, Ap							
462	29	49.2	66	4	US-09-270-767-58777	Sequence 58777, A							
463	29	49.2	68	4	US-09-248-796A-21682	Sequence 21682, A							
464	29	49.2	72	4	US-09-248-796A-23186	Sequence 23186, A							
465	29	49.2	73	4	US-09-248-796A-17432	Sequence 17432, A							

ALIGNMENTS

RESULT 1
US-09-360-376-55
; Sequence 55, Application US/09360376
; Patent No. 6495739
; GENERAL INFORMATION:
; APPLICANT: Lasner, Michael
; APPLICANT: Ruesinsky, Diane
; TITLE OF INVENTION: PLANT PHOSPHATIDIC ACID PHOSPHATASES
; FILE REFERENCE: 17026/01/US
; CURRENT APPLICATION NUMBER: US/09/360,376
; CURRENT FILING DATE: 1999-07-23
; PRIOR APPLICATION NUMBER: US 09/122,315
; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 55
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-360-376-55
Query Match 100.0%; Score 59; DB 4; Length 233;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NYRCRGDSK 10
DB 178 NYRCRGDSK 187
RESULT 2
US-09-489-039A-12030
; Sequence 12030, Application US/09489039A


```
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12030
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12030

Query Match 67.8%; Score 40; DB 4; Length 424;
Best Local Similarity 85.7%; Pred. No. 54;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDD 8
DB 63 FRCRGDD 69

RESULT 3
US-08-311-731A-4
; Sequence 4, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; APPLICANT: MAO, JEN-I
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P. C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,731A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: GATES, EDWARD R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: C0044/7125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/720-3500
; TELEFAX: 617/720-2441
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 705 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: MYCOBACTERIUM TUBERCULOSIS
US-08-311-731A-4

Query Match 67.8%; Score 40; DB 4; Length 705;
Best Local Similarity 66.7%; Pred. No. 88;
```

```
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 YRCRGDDSK 10
DB 383 YRCRADDTTE 391

RESULT 4
US-09-270-767-41177
; Sequence 41177, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41177
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-41177

Query Match 66.1%; Score 39; DB 4; Length 501;
Best Local Similarity 75.0%; Pred. No. 92;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 YRCRGDDSS 9
DB 485 YRCRGDDSS 492

RESULT 5
US-09-270-767-56393
; Sequence 56393, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 56393
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-56393

Query Match 66.1%; Score 39; DB 4; Length 501;
Best Local Similarity 75.0%; Pred. No. 92;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 YRCRGDDSS 9
DB 485 YRCRGDDSS 492

RESULT 6
US-09-489-039A-11789
; Sequence 11789, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
```

FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 11789
LENGTH: 275
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11789

Query Match 62.7%; Score 37; DB 4; Length 275;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDD 9
Db 32 NWRCGGDS 40

RESULT 7
US-09-489-039A-13164
Sequence 13164, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 13164
LENGTH: 488
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13164

Query Match 62.7%; Score 37; DB 4; Length 488;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RCRGDD 8
Db 102 RCRGDD 107

RESULT 8
US-09-949-016-8237
Sequence 8237, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8237
LENGTH: 416
TYPE: PRT
ORGANISM: Human

US-09-949-016-8237

Query Match 61.0%; Score 36; DB 4; Length 416;
Best Local Similarity 60.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
Db 400 NWRCPGTDSE 409

RESULT 9
US-09-786-240-1
Sequence 1, Application US/09786240
Patent No. 6558935
GENERAL INFORMATION:
APPLICANT: INCYTE PHARMACEUTICALS, INC.
APPLICANT: TANG, Y. Tom
APPLICANT: CORLEY, Neil C.
APPLICANT: GUEGLER, Karl J.
APPLICANT: BAUGHN, Mariah R.
APPLICANT: LAL, Preeti
APPLICANT: YUE, Henry
APPLICANT: HILLMAN, Jennifer L.
APPLICANT: AZIMZAI, Yalda
TITLE OF INVENTION: HUMAN TRANSFERASE PROTEINS
FILE REFERENCE: PF-0592 PCT
CURRENT APPLICATION NUMBER: US/09/786,240
CURRENT FILING DATE: 2002-03-12
PRIOR APPLICATION NUMBER: 09/150,657; unassigned; 09/186,779; unassigned; 60/133,642
PRIOR FILING DATE: 1998-09-10; 1998-09-10; 1998-11-04; 1998-11-04; 1999-05-11
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PERL Program
SEQ ID NO 1
LENGTH: 498
TYPE: PRT
ORGANISM: Homo sapiens
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. 6558935 1632930CD1
US-09-786-240-1

Query Match 61.0%; Score 36; DB 4; Length 498;
Best Local Similarity 60.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
Db 482 NWRCPGTDSE 491

RESULT 10
US-09-198-452A-1290
Sequence 1290, Application US/09198452A
Patent No. 6559294
GENERAL INFORMATION:
APPLICANT: Grifflais, R.
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention
TITLE OF INVENTION: and treatment of infection
FILE REFERENCE: 9710-003-999
CURRENT APPLICATION NUMBER: US/09/198,452A
CURRENT FILING DATE: 1998-11-24
NUMBER OF SEQ ID NOS: 6849
SEQ ID NO 1290
LENGTH: 94
TYPE: PRT
ORGANISM: Chlamydia pneumoniae
US-09-198-452A-1290

Query Match 59.3%; Score 35; DB 4; Length 94;
Best Local Similarity 85.7%; Pred. No. 85;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY      1 NVRCRGD 7
      |||||
Db      13 NRRCRGD 19

RESULT 11
US-10-000-489-78
; Sequence 78, Application US/10000489
; Patent No. 6794363
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US6.DIV
; CURRENT APPLICATION NUMBER: US/10/000,489
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: PCT/IB01/01715
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 78
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: 1..21
US-10-000-489-78

Query Match      59.3%; Score 35; DB 4; Length 269;
Best Local Similarity 66.7%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2 YRCRGDDSK 10
      |||||
Db      183 YRCNGTSSK 191

RESULT 12
US-09-949-016-6121
; Sequence 6121, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6121
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6121

Query Match      59.3%; Score 35; DB 4; Length 269;
Best Local Similarity 66.7%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2 YRCRGDDSK 10
      |||||
Db      183 YRCNGTSSK 191

RESULT 13
US-09-949-016-7261
; Sequence 7261, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7261
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7261

Query Match      59.3%; Score 35; DB 4; Length 276;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2 YRCRGDDSK 10
      |||||
Db      190 YRCNGTSSK 198

RESULT 14
US-08-985-950-12
; Sequence 12, Application US/08985950
; Patent No. 6140076
; GENERAL INFORMATION:
; APPLICANT: Adema, Gosse Jan
; TITLE OF INVENTION: Isolated Mammalian Monocyte Cell Genes;
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/985,950
; FILING DATE: 05-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/041,279
; FILING DATE: 21-MARCH-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/033,181
```

```
/
/ FILING DATE: 16-DEC-1996
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 60/032,252
/ FILING DATE: 06-DEC-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Ching, Edwin P.
/ REGISTRATION NUMBER: 34,090
/ REFERENCE/DOCKET NUMBER: DX0670K
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (650)852-9196
/ TELEFAX: (650)496-1204
/ INFORMATION FOR SEQ ID NO: 12:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 439 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-985-950-12

Query Match 59.3%; Score 35; DB 3; Length 439;
Best Local Similarity 60.0%; Pred. No. 3.7e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
Db :||| |||:
194 SYRCYGDSR 203

RESULT 15
US-09-310-463-8
/ Sequence 8, Application US/09310463A
/ Patent No. 6384203
/ GENERAL INFORMATION:
/ APPLICANT: Cosman, David J.
/ APPLICANT: Anderson, Dirk M.
/ APPLICANT: Borges, Luis
/ TITLE OF INVENTION: Family of Immunoregulators Designated Leukocyte Immunoglobulin-
/ TITLE OF INVENTION: Like Receptors (LIR)
/ FILE REFERENCE: 2624-A
/ CURRENT APPLICATION NUMBER: US/09/310,463A
/ CURRENT FILING DATE: 1999-05-12
/ EARLIER APPLICATION NUMBER: 08/842,248
/ EARLIER FILING DATE: 1997-04-24
/ NUMBER OF SEQ ID NOS: 39
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO: 8
/ LENGTH: 439
/ TYPE: PRT
/ ORGANISM: human
/ US-09-310-463-8

Query Match 59.3%; Score 35; DB 3; Length 439;
Best Local Similarity 60.0%; Pred. No. 3.7e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
Db :||| |||:
194 SYRCYGDSR 203

RESULT 16
US-08-842-248A-8
/ Sequence 8, Application US/08842248A
/ Patent No. 6448035
/ GENERAL INFORMATION:
/ APPLICANT: Cosman, David J.
/ TITLE OF INVENTION: Family of Immunoregulators Designated
/ TITLE OF INVENTION: Leukocyte Immunoglobulin-Like Receptors (LIR)
/ NUMBER OF SEQUENCES: 29
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Janis C. Henry, Immunex Corporation
/ STREET: 51 University Street
/ CITY: Seattle
```

```
/
/ STATE: WA
/ COUNTRY: US
/ ZIP: 98101
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM/PC Compatible
/ OPERATING SYSTEM: Microsoft Word 7.0
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/842,248A
/ FILING DATE: April 24, 1997
/ CLASSIFICATION: 536
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Henry, Janis C.
/ REGISTRATION NUMBER: 34,347
/ REFERENCE/DOCKET NUMBER: 2624
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (206) 587-0430
/ TELEFAX: (206) 233-0644
/ TELEX: 756822
/ INFORMATION FOR SEQ ID NO: 8:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 439 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-842-248A-8

Query Match 59.3%; Score 35; DB 4; Length 439;
Best Local Similarity 60.0%; Pred. No. 3.7e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
Db :||| |||:
194 SYRCYGDSR 203

RESULT 17
US-09-546-049-12
/ Sequence 12, Application US/09546049
/ Patent No. 6479638
/ GENERAL INFORMATION:
/ APPLICANT: Adema, Gosse Jan
/ APPLICANT: Meyvaard, Linde
/ APPLICANT: Gorman, Daniel M.
/ APPLICANT: McClanahan, Terrill K.
/ APPLICANT: Zurawski, Sandra M.
/ APPLICANT: Zurawski, Gerard
/ APPLICANT: Lanier, Lewis L.
/ APPLICANT: Phillips Jr., Joseph H.
/ TITLE OF INVENTION: Isolated Mammalian Monocyte Cell Genes;
/ RELATED REAGENTS
/ NUMBER OF SEQUENCES: 22
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: DNAX Research Institute
/ STREET: 901 California Avenue
/ CITY: Palo Alto
/ STATE: California
/ COUNTRY: USA
/ ZIP: 94304-1104
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/546,049
/ FILING DATE: 10-Apr-2000
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/985,950
/ FILING DATE: 05-DEC-1997
/ APPLICATION NUMBER: US 60/041,279
/ FILING DATE: 21-MARCH-1997
```

; APPLICATION NUMBER: US 60/033,181
 ; FILING DATE: 16-DEC-1996
 ; APPLICATION NUMBER: US 60/032,252
 ; FILING DATE: 06-DEC-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Ching, Edwin P.
 ; REGISTRATION NUMBER: 34,090
 ; REFERENCE/DOCKET NUMBER: DX0670K
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (650)852-9196
 ; TELEFAX: (650)496-1204
 ; INFORMATION FOR SEQ ID NO: 12:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 439 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
 US-09-546-049-12

Query Match 59.3%; Score 35; DB 4; Length 439;
 Best Local Similarity 60.0%; Pred. No. 3.7e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
 DB 194 SYRCGYDSR 203

RESULT 18
 US-09-248-796A-17110
 ; Sequence 17110, Application US/09248796A
 ; Patent No. 6747137
 ; GENERAL INFORMATION:
 ; APPLICANT: Keith Weinstock et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
 ; FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.132
 ; CURRENT APPLICATION NUMBER: US/09/248,796A
 ; PRIOR FILING DATE: 1999-02-12
 ; PRIOR APPLICATION NUMBER: US 60/074,725
 ; PRIOR FILING DATE: 1998-02-13
 ; PRIOR APPLICATION NUMBER: US 60/096,409
 ; PRIOR FILING DATE: 1998-08-13
 ; NUMBER OF SEQ ID NOS: 28208
 ; SEQ ID NO 17110
 ; LENGTH: 503
 ; TYPE: PRT
 ; ORGANISM: Candida albicans

US-09-248-796A-17110

Query Match 59.3%; Score 35; DB 4; Length 503;
 Best Local Similarity 66.7%; Pred. No. 4.2e+02;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 YRCRGDDSK 10
 DB 288 YRCFSDDDK 296

RESULT 19
 US-08-419-652-4
 ; Sequence 4, Application US/08419652
 ; Patent No. 5831007
 ; GENERAL INFORMATION:
 ; APPLICANT: Chua, Anne O
 ; APPLICANT: Gubler, Ulrich A
 ; TITLE OF INVENTION: INTERLEUKIN-12 RECEPTOR
 ; NUMBER OF SEQUENCES: 7
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Hoffmann-La Roche Inc.
 ; STREET: 340 Kingsland Street
 ; CITY: Nutley

; STATE: New Jersey
 ; COUNTRY: United States of America
 ; ZIP: 07110-1199
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: PC compatible
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/419,652
 ; FILING DATE: 11-APR-1995
 ; CLASSIFICATION: 530
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/248,532
 ; FILING DATE: 31-MAY-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/094,713
 ; FILING DATE: 19-JUL-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kass, Alan P
 ; REGISTRATION NUMBER: 32142
 ; REFERENCE/DOCKET NUMBER: CD 9174
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (201) 235-4205
 ; TELEFAX: (201) 235-3500
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 621 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; FEATURE:
 ; NAME/KEY: Region
 ; LOCATION: 1..621
 ; OTHER INFORMATION: /note= "Represents residues 42 to
 ; OTHER INFORMATION: 662 of SEQ ID NO:2."
 US-08-419-652-4

Query Match 59.3%; Score 35; DB 2; Length 621;
 Best Local Similarity 75.0%; Pred. No. 5.2e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 RCRGDDSK 10
 DB 443 RCRDEDSK 450

RESULT 20
 US-08-248-532-3
 ; Sequence 3, Application US/08248532
 ; Patent No. 5536657
 ; GENERAL INFORMATION:
 ; APPLICANT: Chua, Anne O
 ; APPLICANT: Gubler, Ulrich A
 ; TITLE OF INVENTION: INTERLEUKIN-12 RECEPTOR
 ; NUMBER OF SEQUENCES: 3
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Hoffmann-La Roche Inc.
 ; STREET: 340 Kingsland Street
 ; CITY: Nutley
 ; STATE: New Jersey
 ; COUNTRY: United States of America
 ; ZIP: 07110-1199
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/248,532
 ; FILING DATE:
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:

```

; APPLICATION NUMBER: 08/094,713
; FILING DATE: 19-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Kass, Alan P
; REGISTRATION NUMBER: 32142
; REFERENCE/DOCKET NUMBER: CD 8973
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 235-4205
; TELEFAX: (201) 235-3500
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-248-532-3

Query Match 59.3%; Score 35; DB 1; Length 660;
Best Local Similarity 75.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 RCRGDSK 10
DB 484 RCRDEDSK 491

RESULT 21
US-08-419-652-3
; Sequence 3, Application US/08419652
; Patent No. 5831007
; GENERAL INFORMATION:
; APPLICANT: Chua, Anne O
; APPLICANT: Gubler, Ulrich A
; TITLE OF INVENTION: INTERLEUKIN-12 RECEPTOR
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: New Jersey
; COUNTRY: United States of America
; ZIP: 07110-1199
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/419,652
; FILING DATE: 11-APR-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/248,532
; FILING DATE: 31-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/094,713
; FILING DATE: 19-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Kass, Alan P
; REGISTRATION NUMBER: 32142
; REFERENCE/DOCKET NUMBER: CD 9174
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 235-4205
; TELEFAX: (201) 235-3500
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-419-652-3

Query Match 59.3%; Score 35; DB 2; Length 660;

```

```

Best Local Similarity 75.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 RCRGDSK 10
DB 484 RCRDEDSK 491

RESULT 22
US-08-248-532-2
; Sequence 2, Application US/08248532
; Patent No. 5536657
; GENERAL INFORMATION:
; APPLICANT: Chua, Anne O
; APPLICANT: Gubler, Ulrich A
; TITLE OF INVENTION: INTERLEUKIN-12 RECEPTOR
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: New Jersey
; COUNTRY: United States of America
; ZIP: 07110-1199
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/248,532
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/094,713
; FILING DATE: 19-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Kass, Alan P
; REGISTRATION NUMBER: 32142
; REFERENCE/DOCKET NUMBER: CD 8973
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 235-4205
; TELEFAX: (201) 235-3500
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 662 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Region
; LOCATION: 1..23
; OTHER INFORMATION: /note= "N-terminal signal peptide
; (1..20 or 23 or 24)"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 541..570
; OTHER INFORMATION: /note= "transmembrane region"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 571..662
; OTHER INFORMATION: /note= "cytoplasmic tail region"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 577..584
; OTHER INFORMATION: /note= "conserved area of cytoplasmic
; tail region"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 618..629
; OTHER INFORMATION: /note= "conserved area of cytoplasmic
; tail region"
; FEATURE:

```

```

; NAME/KEY: Region
; LOCATION: 52..64
; OTHER INFORMATION: /note= "sequence motif of cytokine
; OTHER INFORMATION: receptor superfamily Cys52..Cys62SW"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 222..226
; OTHER INFORMATION: /note= "cytokine receptor
; OTHER INFORMATION: superfamily motif (W222SKWS)"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 121..123
; OTHER INFORMATION: /note= "N-linked glycosylation
; OTHER INFORMATION: site"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 329..331
; OTHER INFORMATION: /note= "N-linked glycosylation
; OTHER INFORMATION: site"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 346..348
; OTHER INFORMATION: /note= "N-linked glycosylation
; OTHER INFORMATION: site"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 352..354
; OTHER INFORMATION: /note= "N-linked glycosylation
; OTHER INFORMATION: site"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 442..444
; OTHER INFORMATION: /note= "N-linked glycosylation
; OTHER INFORMATION: site"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 456..458
; OTHER INFORMATION: /note= "N-linked glycosylation
; OTHER INFORMATION: site"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 24..540
; OTHER INFORMATION: /note= "Extracellular region"
;
US-08-248-532-2
Query Match 59.3%; Score 35; DB 1; Length 662;
Best Local Similarity 75.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 RCRGDSK 10
DB 484 RCRDESK 491

RESULT 23
US-08-419-652-2
; Sequence 2, Application US/08419652
; Patent No. 5831007
; GENERAL INFORMATION:
; APPLICANT: Chua, Anne O
; APPLICANT: Gubler, Ulrich A
; TITLE OF INVENTION: INTERLEUKIN-12 RECEPTOR
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: New Jersey
; COUNTRY: United States of America
; ZIP: 07110-1199
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: PC compatible
```

```

; OPERATING SYSTEM: MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/419,652
; FILING DATE: 11-APR-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/248,532
; FILING DATE: 31-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/094,713
; FILING DATE: 19-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Kass, Alan P
; REGISTRATION NUMBER: 32142
; REFERENCE/DOCKET NUMBER: CD 9174
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 235-4205
; TELEFAX: (201) 235-3500
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 662 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Region
; LOCATION: 1..23
; OTHER INFORMATION: /note= "N-terminal signal peptide
; OTHER INFORMATION: (1..20 or 23 or 24)"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 541..570
; OTHER INFORMATION: /note= "transmembrane region"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 571..662
; OTHER INFORMATION: /note= "cytoplasmic tail region"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 577..584
; OTHER INFORMATION: /note= "conserved area of cytoplasmic
; OTHER INFORMATION: tail region"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 618..629
; OTHER INFORMATION: /note= "conserved area of cytoplasmic
; OTHER INFORMATION: tail region"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 52..64
; OTHER INFORMATION: /note= "sequence motif of cytokine
; OTHER INFORMATION: receptor superfamily Cys52..Cys62SW"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 222..226
; OTHER INFORMATION: /note= "cytokine receptor
; OTHER INFORMATION: superfamily motif (W222SKWS)"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 121..123
; OTHER INFORMATION: /note= "N-linked glycosylation
; OTHER INFORMATION: site"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 329..331
; OTHER INFORMATION: /note= "N-linked glycosylation
; OTHER INFORMATION: site"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 346..348
; OTHER INFORMATION: /note= "N-linked glycosylation
; OTHER INFORMATION: site"
```

```

;
; FEATURE:
; NAME/KEY: Region
; LOCATION: 352..354
; OTHER INFORMATION: /note= "N-linked glycosylation
; OTHER INFORMATION: site"
;
; FEATURE:
; NAME/KEY: Region
; LOCATION: 442..444
; OTHER INFORMATION: /note= "N-linked glycosylation
; OTHER INFORMATION: site"
;
; FEATURE:
; NAME/KEY: Region
; LOCATION: 456..458
; OTHER INFORMATION: /note= "N-linked glycosylation
; OTHER INFORMATION: site"
;
; NAME/KEY: Region
; LOCATION: 24..540
; OTHER INFORMATION: /note= "Extracellular region"
;
US-08-419-652-2
;
; Query Match 59.3%; Score 35; DB 2; Length 662;
; Best Local Similarity 75.0%; Pred. NO. 5.5e+02;
; Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
;
QY 3 RCRGDDSK 10
||| :|||
Db 484 RCRDEDSK 491

RESULT 24
US-08-685-118-4
; Sequence 4, Application US/08695118
; Patent No. 5840530
; GENERAL INFORMATION:
; APPLICANT: Gubler, Ulrich A
; APPLICANT: Presky, David H
; TITLE OF INVENTION: RECEPTORS FOR HUMAN IL-12
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: NJ
; COUNTRY: USA
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/685,118
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Silverman, Robert A.
; REGISTRATION NUMBER: 35,682
; REFERENCE/DOCKET NUMBER: CD 9195
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 235-2863
; TELEFAX: (201) 235-2363
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 662 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Region
; LOCATION: 1..20
; OTHER INFORMATION: /note= "N-terminal signal peptide
; OTHER INFORMATION: (1..20 or 23 or 24)"

```

```

;
; FEATURE:
; NAME/KEY: Region
; LOCATION: 541..570
; OTHER INFORMATION: /note= "transmembrane region"
;
; FEATURE:
; NAME/KEY: Region
; LOCATION: 571..662
; OTHER INFORMATION: /note= "cytoplasmic tail region"
;
; FEATURE:
; NAME/KEY: Region
; LOCATION: 52..64
; OTHER INFORMATION: /note= "sequence motif of cytokine
; OTHER INFORMATION: receptor superfamily Cys52..Cys62SW"
;
; FEATURE:
; NAME/KEY: Region
; LOCATION: 222..226
; OTHER INFORMATION: /note= "cytokine receptor
; OTHER INFORMATION: superfamily motif (W222SKWS)"
;
; FEATURE:
; NAME/KEY: Region
; LOCATION: 121..123
; OTHER INFORMATION: /note= "N-linked glycosylation
; OTHER INFORMATION: site"
;
; FEATURE:
; NAME/KEY: Region
; LOCATION: 329..331
; OTHER INFORMATION: /note= "N-linked glycosylation
; OTHER INFORMATION: site"
;
; FEATURE:
; NAME/KEY: Region
; LOCATION: 346..348
; OTHER INFORMATION: /note= "N-linked glycosylation
; OTHER INFORMATION: site"
;
; FEATURE:
; NAME/KEY: Region
; LOCATION: 352..354
; OTHER INFORMATION: /note= "N-linked glycosylation
; OTHER INFORMATION: site"
;
; FEATURE:
; NAME/KEY: Region
; LOCATION: 442..444
; OTHER INFORMATION: /note= "N-linked glycosylation
; OTHER INFORMATION: site"
;
; FEATURE:
; NAME/KEY: Region
; LOCATION: 456..458
; OTHER INFORMATION: /note= "N-linked glycosylation
; OTHER INFORMATION: site"
;
; FEATURE:
; NAME/KEY: Region
; LOCATION: 24..540
; OTHER INFORMATION: /note= "Extracellular region"
;
US-08-685-118-4
;
; Query Match 59.3%; Score 35; DB 2; Length 662;
; Best Local Similarity 75.0%; Pred. NO. 5.5e+02;
; Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
;
QY 3 RCRGDDSK 10
||| :|||
Db 484 RCRDEDSK 491

RESULT 25
US-08-915-495-4
; Sequence 4, Application US/08915495
; Patent No. 5852176
; GENERAL INFORMATION:
; APPLICANT: Gubler, Ulrich A
; APPLICANT: Presky, David H
; TITLE OF INVENTION: RECEPTORS FOR HUMAN IL-12
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:

```


ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: NJ
COUNTRY: USA
ZIP: 07110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,495
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/685,118
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Silverman, Robert A.
REGISTRATION NUMBER: 35,682
REFERENCE/DOCKET NUMBER: CD 9195
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201) 235-2863
TELEFAX: (201) 235-2363
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 662 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Region
LOCATION: 1..20
OTHER INFORMATION: /note= "N-terminal signal peptide
OTHER INFORMATION: (1..20 or 23 or 24)"
FEATURE:
NAME/KEY: Region
LOCATION: 541..570
OTHER INFORMATION: /note= "transmembrane region"
FEATURE:
NAME/KEY: Region
LOCATION: 571..662
OTHER INFORMATION: /note= "cytoplasmic tail region"
FEATURE:
NAME/KEY: Region
LOCATION: 52..64
OTHER INFORMATION: /note= "sequence motif of cytokine
OTHER INFORMATION: receptor superfamily Cys52..Cys62SW"
FEATURE:
NAME/KEY: Region
LOCATION: 222..226
OTHER INFORMATION: /note= "cytokine receptor
OTHER INFORMATION: superfamily motif (W222SKWS)"
FEATURE:
NAME/KEY: Region
LOCATION: 121..123
OTHER INFORMATION: /note= "N-linked glycosylation
OTHER INFORMATION: site"
FEATURE:
NAME/KEY: Region
LOCATION: 329..331
OTHER INFORMATION: /note= "N-linked glycosylation
OTHER INFORMATION: site"
FEATURE:
NAME/KEY: Region
LOCATION: 346..348
OTHER INFORMATION: /note= "N-linked glycosylation
OTHER INFORMATION: site"
FEATURE:
NAME/KEY: Region
LOCATION: 352..354
OTHER INFORMATION: /note= "N-linked glycosylation

OTHER INFORMATION: site"
FEATURE:
NAME/KEY: Region
LOCATION: 442..444
OTHER INFORMATION: /note= "N-linked glycosylation
OTHER INFORMATION: site"
FEATURE:
NAME/KEY: Region
LOCATION: 456..458
OTHER INFORMATION: /note= "N-linked glycosylation
OTHER INFORMATION: site"
FEATURE:
NAME/KEY: Region
LOCATION: 24..540
OTHER INFORMATION: /note= "Extracellular region"
US-08-915-495-4
Query Match 59.3%; Score 35; DB 2; Length 662;
Best Local Similarity 75.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 3 RCRGDSK 10
Db 484 RCRDEDSK 491
RESULT 26
US-08-914-520-4
Sequence 4; Application US/08914520
Patent No. 5919903
GENERAL INFORMATION:
APPLICANT: Gubler, Ulrich A
APPLICANT: Presky, David H
TITLE OF INVENTION: RECEPTORS FOR HUMAN IL-12
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: NJ
COUNTRY: USA
ZIP: 07110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/914,520
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/685,118
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Silverman, Robert A.
REGISTRATION NUMBER: 35,682
REFERENCE/DOCKET NUMBER: CD 9195
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201) 235-2863
TELEFAX: (201) 235-2363
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 662 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Region
LOCATION: 1..20
OTHER INFORMATION: /note= "N-terminal signal peptide
OTHER INFORMATION: (1..20 or 23 or 24)"
FEATURE:

```
/ NAME/KEY: Region
/ LOCATION: 541..570 /note= "transmembrane region"
/ OTHER INFORMATION:
/ FEATURE:
/ NAME/KEY: Region
/ LOCATION: 571..662 /note= "cytoplasmic tail region"
/ OTHER INFORMATION:
/ FEATURE:
/ NAME/KEY: Region
/ LOCATION: 52..64
/ OTHER INFORMATION: /note= "sequence motif of cytokine
/ OTHER INFORMATION: receptor superfamily Cys52..Cys62SW"
/ FEATURE:
/ NAME/KEY: Region
/ LOCATION: 222..226
/ OTHER INFORMATION: /note= "cytokine receptor
/ OTHER INFORMATION: superfamily motif (W222SKWS)"
/ FEATURE:
/ NAME/KEY: Region
/ LOCATION: 121..123
/ OTHER INFORMATION: /note= "N-linked glycosylation
/ OTHER INFORMATION: site"
/ FEATURE:
/ NAME/KEY: Region
/ LOCATION: 329..331
/ OTHER INFORMATION: /note= "N-linked glycosylation
/ OTHER INFORMATION: site"
/ FEATURE:
/ NAME/KEY: Region
/ LOCATION: 346..348
/ OTHER INFORMATION: /note= "N-linked glycosylation
/ OTHER INFORMATION: site"
/ FEATURE:
/ NAME/KEY: Region
/ LOCATION: 352..354
/ OTHER INFORMATION: /note= "N-linked glycosylation
/ OTHER INFORMATION: site"
/ FEATURE:
/ NAME/KEY: Region
/ LOCATION: 442..444
/ OTHER INFORMATION: /note= "N-linked glycosylation
/ OTHER INFORMATION: site"
/ FEATURE:
/ NAME/KEY: Region
/ LOCATION: 456..458
/ OTHER INFORMATION: /note= "N-linked glycosylation
/ OTHER INFORMATION: site"
/ FEATURE:
/ NAME/KEY: Region
/ LOCATION: 24..540
/ OTHER INFORMATION: /note= "Extracellular region"
US-08-914-520-4
```

```
Query Match 59.3%; Score 35; DB 2; Length 662;
Best Local Similarity 75.0%; Pred. NO. 5.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 3 RCRGDDSK 10
|||:||||
Db 484 RCRDEDSK 491
```

```
RESULT 27
US-08-789-350-2
; Sequence 2, Application US/08789350
; Patent No. 6046012
; GENERAL INFORMATION:
; APPLICANT: Chizzonite, Richard A
; TITLE OF INVENTION: Antibody to IL-12 Receptor
; FILE REFERENCE: Antibody to IL-12
; CURRENT APPLICATION NUMBER: US/08/789,350
; CURRENT FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 08/248,531
```

```
/ EARLIER FILING DATE: 1994-05-31
/ NUMBER OF SEQ ID NOS: 2
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 2
/ LENGTH: 662
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (1)..(23)
/ OTHER INFORMATION: N-terminal signal peptide (1...20 or 23 or 24)
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (541)..(570)
/ OTHER INFORMATION: transmembrane region
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (571)..(662)
/ OTHER INFORMATION: cytoplasmic tail region
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (52)..(64)
/ OTHER INFORMATION: sequence motif of cytokine receptor superfamily
/ OTHER INFORMATION: Cys52...Cys62SW
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (222)..(226)
/ OTHER INFORMATION: cytokine receptor superfamily motif (W222SKWS)
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (121)..(123)
/ OTHER INFORMATION: N-linked glycosylation site
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (329)..(331)
/ OTHER INFORMATION: N-linked glycosylation site
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (346)..(348)
/ OTHER INFORMATION: N-linked glycosylation site
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (352)..(354)
/ OTHER INFORMATION: N-linked glycosylation site
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (442)..(444)
/ OTHER INFORMATION: N-linked glycosylation site
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (456)..(458)
/ OTHER INFORMATION: N-linked glycosylation site
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (24)..(540)
/ OTHER INFORMATION: Extracellular region
US-08-789-350-2
```

```
Query Match 59.3%; Score 35; DB 3; Length 662;
Best Local Similarity 75.0%; Pred. NO. 5.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 3 RCRGDDSK 10
|||:||||
Db 484 RCRDEDSK 491
```

```
RESULT 28
US-09-949-016-6624
; Sequence 6624, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
```

```
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6624
; LENGTH: 662
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6624

Query Match          59.3%; Score 35; DB 4; Length 662;
Best Local Similarity 75.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      3 RCRGDDSK 10
Db      484 RCRDEDSK 491

RESULT 29
US-09-949-016-10092
; Sequence 10092, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10092
; LENGTH: 683
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10092

Query Match          59.3%; Score 35; DB 4; Length 683;
Best Local Similarity 75.0%; Pred. No. 5.6e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      3 RCRGDDSK 10
Db      505 RCRDEDSK 512

RESULT 30
US-09-252-991A-23746
; Sequence 23746, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
```

```
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23746
; LENGTH: 700
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23746

Query Match          59.3%; Score 35; DB 4; Length 700;
Best Local Similarity 75.0%; Pred. No. 5.8e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2 YRCRGDGS 9
Db      530 YRCRGDLA 537

RESULT 31
US-09-252-991A-18617
; Sequence 18617, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18617
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18617

Query Match          57.6%; Score 34; DB 4; Length 163;
Best Local Similarity 85.7%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      3 RCRGDDDS 9
Db      82 RCRGPDS 88

RESULT 32
US-09-489-039A-13748
; Sequence 13748, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13748
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13748

Query Match          57.6%; Score 34; DB 4; Length 272;
Best Local Similarity 60.0%; Pred. No. 3.4e+02;
```

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NVRGDDSK 10
Db 210 NVRFGDDAK 219

RESULT 33
US-09-601-478-5
; Sequence 5, Application US/09601478
; Patent No. 6403785
; GENERAL INFORMATION:
; APPLICANT: Oetuka Pharmaceutical Co., Ltd.
; TITLE OF INVENTION: Human TSC403 gene and human ING1L gene
; FILE REFERENCE: Q60193
; CURRENT APPLICATION NUMBER: US/09/601,478
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: JP H10-134679
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: JP H10-73234
; PRIOR FILING DATE: 1998-03-05
; PRIOR APPLICATION NUMBER: JP H10-38133
; PRIOR FILING DATE: 1998-02-03
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human embryonic brain cDNA library
US-09-601-478-5

Query Match 57.6%; Score 34; DB 4; Length 280;
Best Local Similarity 62.5%; Pred. No. 3.5e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RCRGDDSK 10
Db 257 KCRGDNEK 264

RESULT 34
US-09-601-478-8
; Sequence 8, Application US/09601478
; Patent No. 6403785
; GENERAL INFORMATION:
; APPLICANT: Oetuka Pharmaceutical Co., Ltd.
; TITLE OF INVENTION: Human TSC403 gene and human ING1L gene
; FILE REFERENCE: Q60193
; CURRENT APPLICATION NUMBER: US/09/601,478
; CURRENT FILING DATE: 2000-08-03
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: JP H10-134679
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: JP H10-73234
; PRIOR FILING DATE: 1998-03-05
; PRIOR APPLICATION NUMBER: JP H10-38133
; PRIOR FILING DATE: 1998-02-03
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: human embryonic brain cDNA library
US-09-601-478-8

Query Match 57.6%; Score 34; DB 4; Length 280;
Best Local Similarity 62.5%; Pred. No. 3.5e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RCRGDDSK 10
:||||:|

Db 257 KCRGDNEK 264

RESULT 35
US-09-442-013-13
; Sequence 13, Application US/09442013
; Patent No. 6737232
; GENERAL INFORMATION:
; APPLICANT: Lou, Ying
; APPLICANT: Xu, Xiang
; APPLICANT: Leo, Cindy
; APPLICANT: Huang, Betty
; APPLICANT: Shen, Mary
; TITLE OF INVENTION: NOVEL IAPS ASSOCIATED CELL CYCLE PROTEINS, COMPOSITIONS
; FILE REFERENCE: A-68289/DJB/RMS/DAV
; CURRENT APPLICATION NUMBER: US/09/442,013
; CURRENT FILING DATE: 1999-11-17
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-442-013-13

Query Match 57.6%; Score 34; DB 4; Length 280;
Best Local Similarity 62.5%; Pred. No. 3.5e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RCRGDDSK 10
:||||:|

Db 257 KCRGDNEK 264

RESULT 36
US-09-513-365A-1
; Sequence 1, Application US/09513365A
; Patent No. 6790948
; GENERAL INFORMATION:
; APPLICANT: Harris, Curtis C
; APPLICANT: Nagashima, Makoto
; APPLICANT: Government of United States as represented by the Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: New Tumor Suppressor Gene P33ING2
; FILE REFERENCE: 015280-376100US
; CURRENT APPLICATION NUMBER: US/09/513,365A
; CURRENT FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: US 60/121,891
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: p33ING2
; OTHER INFORMATION: polypeptide sequence
US-09-513-365A-1

Query Match 57.6%; Score 34; DB 4; Length 280;
Best Local Similarity 62.5%; Pred. No. 3.5e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RCRGDDSK 10
:||||:|

Db 257 KCRGDNEK 264

RESULT 37
US-09-513-365A-6
; Sequence 6, Application US/09513365A

```
; Patent No. 6790948
; GENERAL INFORMATION:
; APPLICANT: Harris, Curtis C
; APPLICANT: Nagashima, Makoto
; APPLICANT: Government of United States as represented by the Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: New Tumor Suppressor Gene P31NG2
; FILE REFERENCE: 015280-376100US
; CURRENT APPLICATION NUMBER: US/09/513,365A
; CURRENT FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: US 60/121,891
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: missense
; OTHER INFORMATION: p31NG2 sequence - Arg 153 to Ser
US-09-513-365A-6

Query Match          57.6%; Score 34; DB 4; Length 280;
Best Local Similarity 62.5%; Pred. No. 3.5e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      3 RCRGDDSK 10
DB      257 KCRGDNEK 264
          :||||:
          :||||:

RESULT 38
US-09-902-540-15729
; Sequence 15729, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 15729
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-15729

Query Match          57.6%; Score 34; DB 4; Length 289;
Best Local Similarity 85.7%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      3 RCRGDDDS 9
DB      85 RCPGDDDS 91
          |||
          |||

RESULT 39
US-09-583-110-4948
; Sequence 4948, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
```

```
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 4948
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-4948

Query Match          57.6%; Score 34; DB 4; Length 316;
Best Local Similarity 30.8%; Pred. No. 3.9e+02;
Matches 8; Conservative 2; Mismatches 0; Indels 16; Gaps 1;

QY      1 MYRCR-----GDDSK 10
DB      229 NFRCRSNTFWETIHNKILMGDDSR 254
          :|||
          :|||

RESULT 40
US-09-107-433-5053
; Sequence 5053, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5053:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 321 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...321
; SEQUENCE DESCRIPTION: SEQ ID NO: 5053:
```

US-09-107-433-5053

Query Match 57.6%; Score 34; DB 4; Length 321;
Best Local Similarity 30.8%; Pred. No. 4e+02;
Matches 8; Conservative 2; Mismatches 0; Indels 16; Gaps 1;

QY 1 NYRCR-----GDGSK 10
:||||
Db 234 NFRCRSNSTFWETIHNKEILMGDDSR 259

RESULT 41

US-09-252-991A-30671
; Sequence 30671, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30671
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30671

Query Match 57.6%; Score 34; DB 4; Length 328;
Best Local Similarity 62.5%; Pred. No. 4.1e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RCRGDSK 10
|||||
Db 93 RCRGSDAR 100

RESULT 42

US-09-248-796A-21669
; Sequence 21669, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 21669
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-21669

Query Match 57.6%; Score 34; DB 4; Length 357;
Best Local Similarity 62.5%; Pred. No. 4.4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDD 8
:||||
Db 241 SYVCKGDD 248

RESULT 43

US-09-627-376-10
; Sequence 10, Application US/09627376
; Patent No. 6342385
; GENERAL INFORMATION:
; APPLICANT: Qi, Fengxia
; TITLE OF INVENTION: MUTACIN I BIOSYNTHESIS GENES AND PROTEINS
; FILE REFERENCE: UAB-17402/22
; CURRENT APPLICATION NUMBER: US/09/627,376
; CURRENT FILING DATE: 2001-05-30
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-627-376-10

Query Match 57.6%; Score 34; DB 3; Length 447;
Best Local Similarity 60.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
|||:||||
Db 410 NYHLQGDSSK 419

RESULT 44

US-10-047-676B-10
; Sequence 10, Application US/10047676B
; Patent No. 6699970
; GENERAL INFORMATION:
; APPLICANT: Qi, Fengxia
; APPLICANT: Caufield, Page W.
; TITLE OF INVENTION: MUTACIN I BIOSYNTHESIS GENES AND PROTEINS
; FILE REFERENCE: UAB-17403/22
; CURRENT APPLICATION NUMBER: US/10/047,676B
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 09/627,376
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-10-047-676B-10

Query Match 57.6%; Score 34; DB 4; Length 447;
Best Local Similarity 60.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
|||:||||
Db 410 NYHLQGDSSK 419

RESULT 45

US-09-107-149-19
; Sequence 19, Application US/09107149
; Patent No. 6274363
; GENERAL INFORMATION:
; APPLICANT: Leung, David W.
; APPLICANT: Tompkins, Christopher K.
; TITLE OF INVENTION: PHOSPHATIDYLCHOLINE PHOSPHOLIPASE D
; FILE REFERENCE: 077319/0144
; CURRENT APPLICATION NUMBER: US/09/107,149
; CURRENT FILING DATE: 1998-06-30
; EARLIER APPLICATION NUMBER: 08/768,147
; EARLIER FILING DATE: 1996-12-17
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0

```
; SEQ ID NO 19
; LENGTH: 971
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-107-149-19

Query Match          57.6%; Score 34; DB 3; Length 971;
Best Local Similarity 63.8%; Pred. No. 1.2e+03;
Matches 7; Conservative 2; Mismatches 0; Indels 2; Gaps 1;

QY      1 NVR--CRGDDS 9
      ||| |||::|
Db      846 NYRTMCRGENS 856

RESULT 46
US-08-968-752B-4
; Sequence 4, Application US/08968752B
; Patent No. 6043073
; GENERAL INFORMATION:
; APPLICANT: Frohman, Michael A.
; APPLICANT: Morris, Andrew
; TITLE OF INVENTION: No. 6043073el Phospholipase D Polypeptide and
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ONYX Pharmaceuticals, Inc.
; STREET: 3031 Research Drive
; CITY: Richmond
; STATE: California
; COUNTRY: USA
; ZIP: 94806
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/968,752B
; FILING DATE: 13-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/025,469
; FILING DATE: 05-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Giotta, Gregory J.
; REGISTRATION NUMBER: 32,028
; REFERENCE/DOCKET NUMBER: ONYX2004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-222-9700
; TELEFAX: 510-222-9758
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1036 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-968-752B-4

Query Match          57.6%; Score 34; DB 3; Length 1036;
Best Local Similarity 63.8%; Pred. No. 1.2e+03;
Matches 7; Conservative 2; Mismatches 0; Indels 2; Gaps 1;

QY      1 NVR--CRGDDS 9
      ||| |||::|
Db      808 NYRTMCRGENS 818

RESULT 47
US-09-536-224-4
; Sequence 4, Application US/09536224
; Patent No. 6379665
; GENERAL INFORMATION:
```

```
; APPLICANT: Frohman, Michael A.
; APPLICANT: Morris, Andrew
; TITLE OF INVENTION: No. 6379665el Phospholipase D Polypeptide and
; TITLE OF INVENTION: DNA Sequences
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ONYX Pharmaceuticals, Inc.
; STREET: 3031 Research Drive
; CITY: Richmond
; STATE: California
; COUNTRY: USA
; ZIP: 94806
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/536,224
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/968,752
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Giotta, Gregory J.
; REGISTRATION NUMBER: 32,028
; REFERENCE/DOCKET NUMBER: ONYX2004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-222-9700
; TELEFAX: 510-222-9758
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1036 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-536-224-4

Query Match          57.6%; Score 34; DB 3; Length 1036;
Best Local Similarity 63.6%; Pred. No. 1.2e+03;
Matches 7; Conservative 2; Mismatches 0; Indels 2; Gaps 1;

QY      1 NVR--CRGDDS 9
      ||| |||::|
Db      808 NYRTMCRGENS 818

RESULT 48
US-08-768-147B-2
; Sequence 2, Application US/08768147B
; Patent No. 5859222
; GENERAL INFORMATION:
; APPLICANT: Leung, David W.
; APPLICANT: Tompkins, Christopher
; TITLE OF INVENTION: PHOSPHATIDYLCHOLINE PHOSPHOLIPASE D
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cell Therapeutics, Inc.
; STREET: 200 Elliott Avenue West, Suite 400
; CITY: Seattle
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 98119
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" disk, 1.44Mb, double side, high density
; COMPUTER: Hewlett Packard-IBM Compatible
; OPERATING SYSTEM: MS-DOS Version 6
; SOFTWARE: WORD for WINDOWS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/768,147B
; FILING DATE: 16-Dec-1996
; CLASSIFICATION: 435
```

; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: Provisional US60008768
 ; FILING DATE: 15-Dec-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Faciszewski, Stephen
 ; REGISTRATION NUMBER: 36,131,
 ; REFERENCE/DOCKET NUMBER: 1802A
 ; TELEPHONE: (206)282-7100
 ; TELEFAX: (206)284-6206
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1074
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: polypeptide
 ; HYPOTHETICAL: no
 ; ANTI-SENSE: no
 ; FRAGMENT TYPE:
 ; ORIGINAL SOURCE:
 ; ORGANISM: homo sapien
 ; STRAIN:
 ; INDIVIDUAL ISOLATE:
 ; DEVELOPMENTAL STAGE:
 ; HAPLOTYPE:
 ; TISSUE TYPE:
 ; CELL TYPE:
 ; CELL LINE:
 ; ORGANELLE:
 ; IMMEDIATE SOURCE:
 ; LIBRARY:
 ; CLONE:
 ; POSITION IN GENOME:
 ; CHROMOSOME/SEGMENT:
 ; MAP POSITION:
 ; UNITS:
 ; FEATURE:
 ; NAME/KEY: hPLD1
 ; LOCATION:
 ; IDENTIFICATION METHOD:
 ; OTHER INFORMATION:
 ; PUBLICATION INFORMATION:
 ; AUTHORS:
 ; TITLE:
 ; JOURNAL:
 ; VOLUME:
 ; ISSUE:
 ; PAGES:
 ; DATE:
 ; DOCUMENT NUMBER:
 ; FILING DATE:
 ; PUBLICATION DATE:
 ; RELEVANT RESIDUES IN SEQ ID NO: 2:
 US-08-768-147B-2

Query Match 57.6%; Score 34; DB 2; Length 1074;
 Best Local Similarity 63.6%; Pred. No. 1.3e+03;
 Matches 7; Conservative 2; Mismatches 0; Indels 2; Gaps 1;

QY 1 NVR--CRGDDS 9
 ||| |||:
 Db 846 NYRTMCRGENS 856

RESULT 49
 US-08-968-752B-2
 ; Sequence 2, Application US/08968752B
 ; Patent No. 6043073
 ; GENERAL INFORMATION:
 ; APPLICANT: Frohman, Michael A.
 ; APPLICANT: Morris, Andrew
 ; TITLE OF INVENTION: No. 6043073el Phospholipase D Polypeptide and

; TITLE OF INVENTION: DNA Sequences
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: ONYX Pharmaceuticals, Inc.
 ; STREET: 3031 Research Drive
 ; CITY: Richmond
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94806
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/968,752B
 ; FILING DATE: 13-AUG-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/025,469
 ; FILING DATE: 05-SEP-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Giotta, Gregory J.
 ; REGISTRATION NUMBER: 32,028
 ; REFERENCE/DOCKET NUMBER: ONYX2004
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 510-222-9700
 ; TELEFAX: 510-222-9758
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1074 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-968-752B-2

Query Match 57.6%; Score 34; DB 3; Length 1074;
 Best Local Similarity 63.6%; Pred. No. 1.3e+03;
 Matches 7; Conservative 2; Mismatches 0; Indels 2; Gaps 1;

QY 1 NVR--CRGDDS 9
 ||| |||:
 Db 846 NYRTMCRGENS 856

RESULT 50
 US-09-107-149-3
 ; Sequence 3, Application US/09107149
 ; Patent No. 6274363
 ; GENERAL INFORMATION:
 ; APPLICANT: Leung, David W.
 ; APPLICANT: Tompkins, Christopher K.
 ; TITLE OF INVENTION: PHOSPHATIDYLCHOLINE PHOSPHOLIPASE D
 ; FILE REFERENCE: 077319/0144
 ; CURRENT APPLICATION NUMBER: US/09/107,149
 ; CURRENT FILING DATE: 1998-06-30
 ; EARLIER APPLICATION NUMBER: 08/768,147
 ; EARLIER FILING DATE: 1996-12-17
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 1074
 ; TYPE: PPT
 ; ORGANISM: Homo sapiens
 ; US-09-107-149-3

Query Match 57.6%; Score 34; DB 3; Length 1074;
 Best Local Similarity 63.6%; Pred. No. 1.3e+03;
 Matches 7; Conservative 2; Mismatches 0; Indels 2; Gaps 1;

QY 1 NVR--CRGDDS 9
 ||| |||:
 Db 846 NYRTMCRGENS 856

Search completed: September 7, 2005, 19:58:08
Job time : 18.5581 secs

°



GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2005, 19:37:54 ; Search time 72.093 Seconds
(without alignments)
53.647 Million cell updates/sec

Title: US-10-812-238B-20

Perfect score: 59

Sequence: 1 NYRCRGDDSK 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database : A_Geneseq_16Dec04:*

1: Geneseq19808:*

2: Geneseq19908:*

3: Geneseq20008:*

4: Geneseq20018:*

5: Geneseq20028:*

6: Geneseq20038a:*

7: Geneseq20038b:*

8: Geneseq20048:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59	100.0	311	2	Aaw79286 Human pro
2	59	100.0	311	7	Ade60703 Human pro
3	59	100.0	311	8	Adm72103 Human tau
4	59	100.0	311	8	Adn04415 Antipsori
5	59	100.0	311	8	Adp24274 PRO polyp
6	55	93.2	312	5	Abb99112 Rat phosp
7	55	93.2	312	7	Ade60701 Rat Prote
8	41	69.5	152	4	Abg23434 Novel hum
9	41	69.5	237	4	Aau28253 Novel hum
10	41	69.5	237	8	Adsi2199 Human the
11	41	69.5	421	8	Adj50947 Human nov
12	41	69.5	424	8	Adj50945 Human nov
13	41	69.5	428	6	Ada83857 Human POM
14	41	69.5	481	8	Adr09234 Human pro
15	41	69.5	481	8	Adsi10966 Human the
16	41	69.5	489	6	Adm04303 Human pro
17	40	67.8	400	6	Abu31582 Protein e
18	40	67.8	400	6	Abu45262 Protein e
19	40	67.8	400	6	Abu48051 Protein e
20	40	67.8	400	6	Abu15369 Protein e
21	40	67.8	424	7	Ab065513 Klebsiell
22	40	67.8	705	7	Abd74255 Mycobacte
23	40	67.8	1304	6	Abp72189 Plasmodiu
24	39	66.1	451	4	Abb67864 Drosophil
25	39	66.1	1299	6	Abp72191 Plasmodiu

26	38	64.4	822	2	Aaw99087	Aaw99087 Human ser
27	38	64.4	875	8	Aaw83361	Aaw83361 Human neu
28	38	64.4	875	8	Adq21226	Adq21226 Human sof
29	38	64.4	875	5	Adq89082	Adq89082 Human uro
30	37	62.7	77	5	Abp42738	Abp42738 Human ova
31	37	62.7	92	4	Aau61299	Aau61299 Propionib
32	37	62.7	92	6	Abm57818	Abm57818 Propionib
33	37	62.7	113	4	Aam91935	Aam91935 Human dig
34	37	62.7	113	4	Aam38569	Aam38569 Human col
35	37	62.7	113	5	Abg97621	Abg97621 Human col
36	37	62.7	113	7	Adb92877	Adb92877 Human col
37	37	62.7	140	4	Abb11487	Abb11487 Human sec
38	37	62.7	275	7	Ab065272	Ab065272 Klebsiell
39	37	62.7	488	7	Ab066647	Ab066647 Klebsiell
40	37	62.7	1269	7	ADD46237	ADD46237 Rat Prote
41	37	62.7	1272	7	ADD46239	ADD46239 Human Pro
42	36	61.0	59	4	Aau49621	Aau49621 Propionib
43	36	61.0	59	6	Abm46140	Abm46140 Propionib
44	36	61.0	405	6	Abm69291	Abm69291 Phototrab
45	36	61.0	466	8	Adn20343	Adn20343 Bacterial
46	36	61.0	470	8	Adn20289	Adn20289 Bacterial
47	36	61.0	485	7	Adm04173	Adm04173 Human pro
48	36	61.0	498	3	Aay79209	Aay79209 Human tra
49	36	61.0	498	7	Adn95772	Adn95772 Human BEC
50	36	61.0	498	8	Adp23807	Adp23807 PRO polyp
51	36	61.0	1365	6	Abp72194	Abp72194 Plasmodiu
52	36	61.0	1812	6	Abp70268	Abp70268 Amino aci
53	35	59.3	20	5	Abp56183	Abp56183 Chimeric
54	35	59.3	69	4	Aau44985	Aau44985 Propionib
55	35	59.3	69	6	Abm41504	Abm41504 Propionib
56	35	59.3	94	2	Aay35872	Aay35872 Chlamydia
57	35	59.3	181	4	Abb67977	Abb67977 Drosophil
58	35	59.3	269	2	Aar60558	Aar60558 Human bas
59	35	59.3	269	2	Aay39464	Aay39464 CD147 pro
60	35	59.3	269	5	Aau79095	Aau79095 Human EMM
61	35	59.3	269	6	Abp97416	Abp97416 Human HAB
62	35	59.3	269	6	Abu05047	Abu05047 Human exp
63	35	59.3	269	6	Abu05036	Abu05036 Human exp
64	35	59.3	269	6	Abu05042	Abu05042 Human exp
65	35	59.3	269	6	Abu05041	Abu05041 Human exp
66	35	59.3	269	6	Abu05037	Abu05037 Human exp
67	35	59.3	269	6	Abu05038	Abu05038 Human exp
68	35	59.3	269	6	Abu05043	Abu05043 Human exp
69	35	59.3	269	6	Abu05045	Abu05045 Human exp
70	35	59.3	269	6	Abu05039	Abu05039 Human exp
71	35	59.3	269	6	Abu05040	Abu05040 Human exp
72	35	59.3	269	6	Abu05046	Abu05046 Human exp
73	35	59.3	269	6	AbR48491	AbR48491 Human bas
74	35	59.3	269	7	Adj68581	Adj68581 Human hea
75	35	59.3	269	8	Adi82502	Adi82502 Human mod
76	35	59.3	269	8	Abm81976	Abm81976 Tumour-as
77	35	59.3	269	8	ADR68491	ADR68491 Apoptosis
78	35	59.3	276	3	Aab57003	Aab57003 Human pro
79	35	59.3	276	4	Abu05044	Abu05044 Human exp
80	35	59.3	283	4	Abb68798	Abb68798 Drosophil
81	35	59.3	301	8	ADS26705	ADS26705 Bacterial
82	35	59.3	301	8	ADS27091	ADS27091 Bacterial
83	35	59.3	302	8	ADS26338	ADS26338 Bacterial
84	35	59.3	385	5	ABG34071	ABG34071 Human PRO
85	35	59.3	385	5	ADA011352	ADA011352 Human PRO
86	35	59.3	385	6	ADA43781	ADA43781 Human sec
87	35	59.3	385	6	ADA43549	ADA43549 Human sec
88	35	59.3	385	6	ADA01224	ADA01224 Human PRO
89	35	59.3	385	7	ADA01108	ADA01108 Human sec
90	35	59.3	385	7	ADA43665	ADA43665 Human sec
91	35	59.3	385	7	ADA06927	ADA06927 Human PRO
92	35	59.3	385	7	ADA08415	ADA08415 Novel hum
93	35	59.3	385	7	ADb99708	ADb99708 Human PRO
94	35	59.3	385	7	ADb86991	ADb86991 Human PRO
95	35	59.3	385	7	ADb66146	ADb66146 Human sec
96	35	59.3	385	7	ADb99824	ADb99824 Human PRO
97	35	59.3	385	7	ADb99479	ADb99479 Novel hum
98	35	59.3	385	7	ADb66030	ADb66030 Human sec

99	35	59.3	385	7	ADC23428	Adc23428	Human	tra	172	35	59.3	422	4	AAE09446	Aae09446	Human	sbg
100	35	59.3	385	7	ADC26121	Adc26121	Human	PRO	173	35	59.3	427	2	AAW69231	Aaw69231	Human	pro
101	35	59.3	385	7	ADC04948	Adc04948	Human	PRO	174	35	59.3	427	2	ADG46570	Adg46570	Human	fc
102	35	59.3	385	7	ADE11254	Adel1254	Human	PRO	175	35	59.3	427	8	ADR28644	Adr28644	Human	fc
103	35	59.3	385	7	ADD88185	AdD88185	Human	PRO	176	35	59.3	432	2	ADM04721	Adm04721	Human	pro
104	35	59.3	385	7	ADD95480	AdD95480	Human	sec	177	35	59.3	432	2	AAW82546	Aaw82546	Human	LIR
105	35	59.3	385	7	ADE06410	AdE06410	Human	PRO	178	35	59.3	439	2	AAW53464	Aaw53464	Human	gp4
106	35	59.3	385	7	ADE38185	AdE38185	Human	PRO	179	35	59.3	439	2	AAW62777	Aaw62777	Protein	e
107	35	59.3	385	7	ADD88301	AdD88301	Human	PRO	180	35	59.3	439	3	ABO41171	AbO41171	Leukocyte	
108	35	59.3	385	7	ADD90882	AdD90882	Human	sec	181	35	59.3	439	8	ADO24389	AdO24389	Human	PRO
109	35	59.3	385	7	ADF99437	AdF99437	Human	sec	182	35	59.3	439	8	ADP23998	AdP23998	PRO	polyP
110	35	59.3	385	7	ADG06530	AdG06530	Human	PRO	183	35	59.3	453	8	ADN23929	AdN23929	Bacterial	
111	35	59.3	385	7	ADG05481	AdG05481	Human	PRO	184	35	59.3	453	8	ADN23928	AdN23928	Bacterial	
112	35	59.3	385	7	ADG82482	AdG82482	Human	PRO	185	35	59.3	453	8	ADN23930	AdN23930	Bacterial	
113	35	59.3	385	8	ADE51735	AdE51735	Human	sec	186	35	59.3	473	4	ABG27154	AbG27154	Novel	hum
114	35	59.3	385	8	ADE51881	AdE51881	Human	sec	187	35	59.3	478	4	ABG27155	AbG27155	Novel	hum
115	35	59.3	385	8	ADE37709	AdE37709	Human	sec	188	35	59.3	552	4	AAE09447	Aae09447	Human	sbg
116	35	59.3	385	8	ADE37593	AdE37593	Human	sec	189	35	59.3	575	7	ADC13555	AdC13555	Human	NOV
117	35	59.3	385	8	ADD95364	AdD95364	Human	sec	190	35	59.3	614	8	ADR99920	AdR99920	Immune	Re
118	35	59.3	385	8	ADP38064	AdP38064	Human	PRO	191	35	59.3	615	8	ADQ39859	AdQ39859	Human	myo
119	35	59.3	385	8	ADE76153	AdE76153	Human	PRO	192	35	59.3	616	8	ADR99921	AdR99921	Immune	Re
120	35	59.3	385	8	ADP39476	AdP39476	Human	PRO	193	35	59.3	656	5	ABB93180	AbB93180	Herbicide	
121	35	59.3	385	8	ADE04280	AdE04280	Human	PRO	194	35	59.3	660	2	AAW69633	Aaw69633	Human	int
122	35	59.3	385	8	ADE39877	AdE39877	Human	PRO	195	35	59.3	662	2	AAW12772	Aaw12772	Human	int
123	35	59.3	385	8	ADE19742	AdE19742	Human	PRO	196	35	59.3	662	2	AAW12772	Aaw12772	Human	int
124	35	59.3	385	8	ADE77320	AdE77320	Human	sec	197	35	59.3	662	3	AAV83950	Aav83950	Human	int
125	35	59.3	385	8	ADE65428	AdE65428	Human	PRO	198	35	59.3	662	7	ADE39706	AdE39706	Human	int
126	35	59.3	385	8	ADE76037	AdE76037	Human	PRO	199	35	59.3	662	8	ADK90610	AdK90610	Human	IL-
127	35	59.3	385	8	ADE37948	AdE37948	Human	PRO	200	35	59.3	662	8	ADM33833	AdM33833	Human	IL-
128	35	59.3	385	8	ADE64558	AdE64558	Human	PRO	201	35	59.3	662	8	ADQ39861	AdQ39861	Human	myo
129	35	59.3	385	8	ADE38893	AdE38893	Human	PRO	202	35	59.3	662	8	ADQ39860	AdQ39860	Human	myo
130	35	59.3	385	8	ADE51967	AdE51967	Human	sec	203	35	59.3	694	8	ADO26838	AdO26838	Human	rec
131	35	59.3	385	8	ADP90998	AdP90998	Human	sec	204	35	59.3	700	7	ABO75000	AbO75000	Pseudomon	
132	35	59.3	385	8	ADE38777	AdE38777	Human	PRO	205	35	59.3	702	8	ADQ39863	AdQ39863	Human	myo
133	35	59.3	385	8	ADE37477	AdE37477	Human	sec	206	35	59.3	754	8	ADQ39862	AdQ39862	Human	myo
134	35	59.3	385	8	ADE06294	AdE06294	Human	PRO	207	34	57.6	115	4	AAU48996	Aau48996	Propionib	
135	35	59.3	385	8	ADD90153	AdD90153	Human	sec	208	34	57.6	115	6	ABM45515	AbM45515	Propionib	
136	35	59.3	385	8	ADE38661	AdE38661	Human	PRO	209	34	57.6	117	4	AAU79769	Aau79769	Human	pro
137	35	59.3	385	8	ADE39592	AdE39592	Human	PRO	210	34	57.6	121	4	ABG18440	AbG18440	Novel	hum
138	35	59.3	385	8	ADD89197	AdD89197	Human	PRO	211	34	57.6	127	4	AAW95510	Aaw95510	Human	rep
139	35	59.3	385	8	ADD88964	AdD88964	Human	PRO	212	34	57.6	127	4	ABE96193	AbE96193	Human	tes
140	35	59.3	385	8	ADE19858	AdE19858	Human	PRO	213	34	57.6	131	3	AAV75402	Aav75402	Neisseria	
141	35	59.3	385	8	ADE77436	AdE77436	Human	sec	214	34	57.6	131	3	AAV75400	Aav75400	Neisseria	
142	35	59.3	385	8	ADE65312	AdE65312	Human	PRO	215	34	57.6	135	3	ADL92508	AdL92508	Neisseria	
143	35	59.3	385	8	ADP39360	AdP39360	Human	PRO	216	34	57.6	135	8	ADR41823	AdR41823	Versatille	
144	35	59.3	385	8	ADE38545	AdE38545	Human	sec	217	34	57.6	137	7	ADB65039	AdB65039	Human	pro
145	35	59.3	385	8	ADG11098	AdG11098	Human	sec	218	34	57.6	147	4	AAO07776	Aao07776	Human	pol
146	35	59.3	385	8	ADH10982	AdH10982	Human	sec	219	34	57.6	156	7	ADC39079	AdC39079	Novel	hum
147	35	59.3	385	8	ADH31510	AdH31510	Human	PRO	220	34	57.6	163	7	ABO69871	AbO69871	Pseudomon	
148	35	59.3	385	8	ADH38758	AdH38758	Human	sec	221	34	57.6	172	6	ABU26330	AbU26330	Protein	e
149	35	59.3	385	8	ADH29393	AdH29393	Human	sec	222	34	57.6	182	8	ADN46681	AdN46681	Thermococ	
150	35	59.3	385	8	ADH23696	AdH23696	Human	sec	223	34	57.6	235	8	ADR10154	AdR10154	Human	pro
151	35	59.3	385	8	ADH27026	AdH27026	Human	sec	224	34	57.6	240	3	AAV52199	Aav52199	Human	tum
152	35	59.3	385	8	ADH38294	AdH38294	Novel	hum	225	34	57.6	272	7	ABO67231	AbO67231	Klebsiell	
153	35	59.3	385	8	ADH26910	AdH26910	Human	sec	226	34	57.6	273	5	ABB93016	AbB93016	Herbicide	
154	35	59.3	385	8	ADH38178	AdH38178	Novel	hum	227	34	57.6	280	2	AAV29606	Aav29606	Human	ING
155	35	59.3	385	8	ADH38874	AdH38874	Human	sec	228	34	57.6	280	3	AAQ05772	Aaq05772	Arabidops	
156	35	59.3	385	8	ADH23812	AdH23812	Human	sec	229	34	57.6	280	4	AAE06675	Aae06675	Tumour	su
157	35	59.3	385	8	ADH40187	AdH40187	Human	PRO	230	34	57.6	280	8	ADH09685	AdH09685	Human	hos
158	35	59.3	385	8	ADH40072	AdH40072	Human	PRO	231	34	57.6	280	8	ADH09683	AdH09683	Human	hos
159	35	59.3	385	8	ADH31394	AdH31394	Human	PRO	232	34	57.6	280	8	ADH09686	AdH09686	Human	hos
160	35	59.3	385	8	ADH29272	AdH29272	Human	sec	233	34	57.6	280	8	ADH09684	AdH09684	Human	hos
161	35	59.3	385	8	ADH49487	AdH49487	Novel	hum	234	34	57.6	280	8	ADH09687	AdH09687	Human	sup
162	35	59.3	385	8	ADH51951	AdH51951	Novel	hum	235	34	57.6	280	8	ADR43904	AdR43904	Human	sup
163	35	59.3	385	8	ADH49806	AdH49806	Novel	hum	236	34	57.6	280	8	ADR43899	AdR43899	Human	sup
164	35	59.3	385	8	ADH52407	AdH52407	Novel	hum	237	34	57.6	287	3	AAG49721	Aag49721	Arabidops	
165	35	59.3	385	8	ADH52523	AdH52523	Novel	hum	238	34	57.6	288	3	AAG49733	Aag49733	Arabidops	
166	35	59.3	385	8	ADH58520	AdH58520	Novel	hum	239	34	57.6	289	3	AAQ21354	Aaq21354	Arabidops	
167	35	59.3	385	8	ADH51835	AdH51835	Novel	hum	240	34	57.6	299	3	AAQ05771	Aaq05771	Arabidops	
168	35	59.3	385	8	ADH58396	AdH58396	Novel	hum	241	34	57.6	309	5	AAE22290	Aae22290	Rodent	tu
169	35	59.3	385	8	ADH13593	AdH13593	Novel	hum	242	34	57.6	316	8	ADK48433	AdK48433	Streptoco	
170	35	59.3	385	8	ADK00849	AdK00849	Human	PRO	243	34	57.6	321	3	AAQ21353	Aaq21353	Arabidops	
171	35	59.3	385	8	ADL08590	AdL08590	Human	sec	244	34	57.6	321	8	ADR96418	AdR96418	Novel	S.

245	34	57.6	323	8	ADS24081	Adsd24081 Bacterial	318	33	55.9	490	4	AAW40334	Aam40334 Human pol
246	34	57.6	328	7	ABO81925	Abos1925 Pseudomon	319	33	55.9	490	5	ABB84260	Abb84260 Human kat
247	34	57.6	362	8	ADQ15100	Aadq15100 Human can	320	33	55.9	490	5	AAU74352	Aau74352 Human cyt
248	34	57.6	396	4	ABG14749	Abg14749 Novel hum	321	33	55.9	490	7	ADM05650	Adm05650 Human pro
249	34	57.6	397	4	AAU61160	Aau61160 Propionib	322	33	55.9	542	4	ABG17714	Abg17714 Novel hum
250	34	57.6	397	5	ABM57679	Abm57679 Propionib	323	33	55.9	557	3	AAV87913	Aav87913 D. opposi
251	34	57.6	447	5	ABM09643	Abm09643 MutP prot	324	33	55.9	582	2	AAK92256	Aak92256 Neural ce
252	34	57.6	447	6	ABG72541	Abg72541 Streptoco	325	33	55.9	629	8	ADR31370	Adr31370 Stress re
253	34	57.6	645	6	ABM68401	Abm68401 Photorhab	326	33	55.9	636	4	ABG21734	Abg21734 Novel hum
254	34	57.6	801	4	AAU34022	Aau34022 Staphyloc	327	33	55.9	654	8	ADS27333	Ads27333 Bacterial
255	34	57.6	801	6	AAU36926	Aau36926 Staphyloc	328	33	55.9	654	8	ADS26605	Ads26605 Bacterial
256	34	57.6	801	6	ABU16301	Abu16301 Protein e	329	33	55.9	654	8	ADS26980	Ads26980 Bacterial
257	34	57.6	801	6	ABM70889	Abm70889 Staphyloc	330	33	55.9	654	8	ADN26780	Adn26780 Bacterial
258	34	57.6	808	7	ADB64170	Adb64170 Human pro	331	33	55.9	654	8	ADN26504	Adn26504 Bacterial
259	34	57.6	867	7	ADE71288	Ade71288 Novel hum	332	33	55.9	675	8	ADN20157	Adn20157 Bacterial
260	34	57.6	971	3	AAV44717	Aay44717 Human pho	333	33	55.9	678	5	ABP41931	Abp41931 Human ova
261	34	57.6	971	4	AAB85981	Aab85981 Human pho	334	33	55.9	680	8	ADS28720	Ads28720 Bacterial
262	34	57.6	1036	2	AAW53280	Aaw53280 Human pho	335	33	55.9	686	7	ABO69369	Abos69369 Pseudomon
263	34	57.6	1074	2	AAW53279	Aaw53279 Human pho	336	33	55.9	729	8	ADR08965	Adr08965 Human pro
264	34	57.6	1074	4	AAB85976	Aab85976 Human pho	337	33	55.9	730	2	AAR71379	Aar71379 Grasshopp
265	34	57.6	1074	7	ADDE63117	Adde63117 Human pro	338	33	55.9	761	2	AAR922255	Aar922255 Neural ce
266	34	57.6	1074	7	ADD46309	Add46309 Human Pro	339	33	55.9	800	7	ABO60674	Abos60674 Klebsiell
267	34	57.6	1125	4	ABG18927	Abg18927 Novel hum	340	33	55.9	810	8	ADH72220	Adh72220 Human pro
268	34	57.6	4601	4	ABBS9371	Abbs9371 Drosophil	341	33	55.9	825	7	ABO61952	Abos61952 Klebsiell
269	33	55.9	14	2	AAW63598	Aaw63598 Human com	342	33	55.9	837	8	ADI24547	Adi24547 Human mod
270	33	55.9	14	2	AAW63597	Aaw63597 Human com	343	33	55.9	842	4	AAU36357	Aau36357 Pseudomon
271	33	55.9	21	2	AAR93161	Aar93161 Human blo	344	33	55.9	842	6	ABU38620	Abu38620 Protein e
272	33	55.9	21	2	AAW63606	Aaw63606 Human com	345	33	55.9	897	8	ADI45373	Adi45373 Rice isop
273	33	55.9	21	2	AAW63607	Aaw63607 Human com	346	33	55.9	1004	6	ABU54557	Abu54557 Human NOV
274	33	55.9	21	2	AAW63605	Aaw63605 Human com	347	33	55.9	1004	8	ADH72222	Adh72222 Human pro
275	33	55.9	22	8	ADS33747	Adsd33747 cMET-HGF	348	33	55.9	1004	8	ADH72214	Adh72214 Human pro
276	33	55.9	28	2	AAW63609	Aaw63609 Human com	349	33	55.9	1004	8	ADH72224	Adh72224 Human pro
277	33	55.9	40	8	ADP21704	Adp21704 Human IGE	350	33	55.9	1048	6	AAU44506	Aau44506 BAMV coa
278	33	55.9	51	4	AAU51114	Aau51114 Propionib	351	33	55.9	1138	6	ABJ26107	Abj26107 Aspergill
279	33	55.9	51	6	ABM47633	Abm47633 Propionib	352	33	55.9	1202	2	AAU44507	Aau44507 BAMV coa
280	33	55.9	57	4	AAU59258	Aau59258 Propionib	353	33	55.9	1231	6	ABU07436	Abu07436 Protein d
281	33	55.9	57	6	ABM55777	Abm55777 Propionib	354	33	55.9	1286	4	ABG10455	Abg10455 Novel hum
282	33	55.9	65	2	AAW63591	Aaw63591 Human com	355	33	55.9	1907	6	ABU62071	Abu62071 Human ske
283	33	55.9	69	4	AAW92908	Aaw92908 Human dig	356	33	55.9	2647	8	ADP25446	Adp25446 Plasmodiu
284	33	55.9	70	4	AAU45337	Aau45337 Propionib	357	33	55.9	2931	4	ABB68229	Abb68229 Drosophil
285	33	55.9	70	6	ABM41856	Abm41856 Propionib	358	33	55.9	3069	5	AAE20787	Aae20787 Human C3b
286	33	55.9	103	7	ADC97445	Adc97445 E. faeciu	359	33	55.9	3069	5	AAE20900	Aae20900 Human C3b
287	33	55.9	154	3	ADC97445	Adc97445 E. faeciu	360	33	55.9	3100	5	AAE20789	Aae20789 Human C3b
288	33	55.9	155	6	ABM69477	Abm69477 Photorhab	361	33	55.9	3100	5	AAE20901	Aae20901 Human C3b
289	33	55.9	174	7	ADC95088	Adc95088 E. faeciu	362	33	55.9	3389	7	ADJ70480	Adj70480 Human hea
290	33	55.9	180	4	AAU43388	Aau43388 Propionib	363	33	55.9	3567	8	ADH72216	Adh72216 Human pro
291	33	55.9	180	6	ABM39907	Abm39907 Propionib	364	33	55.9	4753	8	ADN22466	Adn22466 Bacterial
292	33	55.9	184	7	ADF13889	Adf13889 Human end	365	32.5	55.1	76	4	AAU42558	Aau42558 Propionib
293	33	55.9	185	6	ABU35961	Abu35961 Protein e	366	32.5	55.1	76	6	ABM39077	Abm39077 Propionib
294	33	55.9	190	5	AAE18631	Aae18631 Equine hi	367	32	54.2	7	2	AAW96393	Aaw96393 RGD cycli
295	33	55.9	191	6	ABU54593	Abu54593 Human NOV	368	32	54.2	7	2	AAW66831	Aaw66831 Peptide u
296	33	55.9	201	2	AAV27061	Aay27061 Recombina	369	32	54.2	7	2	AAW96984	Aaw96984 Peptide u
297	33	55.9	230	4	AAU53393	Aau53393 Propionib	370	32	54.2	9	2	AAW96408	Aaw96408 RGD cycli
298	33	55.9	230	6	ABM49912	Abm49912 Propionib	371	32	54.2	9	2	AAW66840	Aaw66840 Cyclic pe
299	33	55.9	232	4	AAW42120	Aaw42120 Human pol	372	32	54.2	9	2	AAW97003	Aaw97003 Peptide u
300	33	55.9	236	2	AAV27060	Aay27060 Equine Pc	373	32	54.2	10	1	AAW90621	Aaw90621 Peptide u
301	33	55.9	248	8	ADN18186	Adn18186 Bacterial	374	32	54.2	19	2	AAR13390	Aar13390 (59-78) hN
302	33	55.9	253	7	ABO83713	Abos3713 Pseudomon	375	32	54.2	20	2	ADF77514	Adf77514 DAF-7 pro
303	33	55.9	255	2	AAV27058	Aay27058 Equine Pc	376	32	54.2	20	3	AAW606057	Aaw606057 Caenorhab
304	33	55.9	252	2	AAW92254	Aar92254 Neural ce	377	32	54.2	23	6	AAE30217	Aae30217 Human LP2
305	33	55.9	299	7	ABO73111	Abos7111 Pseudomon	378	32	54.2	23	6	AAE30316	Aae30316 Human LP2
306	33	55.9	299	8	ADH72218	Adh72218 Human pro	379	32	54.2	29	5	AAE18630	Aae18630 Human imm
307	33	55.9	349	4	AAW42035	Aaw42035 Human pol	380	32	54.2	35	2	AAE29504	Aar29504 NGF, rat.
308	33	55.9	352	5	ABB91426	Abb91426 Herbicida	381	32	54.2	35	3	ADF77517	Adf77517 DAF-7 pro
309	33	55.9	411	4	ABB61304	Abb61304 Drosophil	382	32	54.2	35	3	AAW606060	Aaw606060 Caenorhab
310	33	55.9	416	2	ADH11311	Adh11311 Vertebrat	383	32	54.2	35	8	ADP21822	Adp21822 Low densi
311	33	55.9	420	4	ABG26076	Abg26076 Novel hum	384	32	54.2	42	2	AAR29473	Aar29473 NGF, rat.
312	33	55.9	460	4	AAW57313	Aaw57313 Human bra	385	32	54.2	42	2	AAR29472	Aar29472 NGF, huma
313	33	55.9	460	4	ABG51397	Abg51397 Human liv	386	32	54.2	43	2	ADF77515	Adf77515 DAF-7 pro
314	33	55.9	460	4	AAW05198	Aam05198 Peptide #	387	32	54.2	43	3	AAW606058	Aaw606058 Caenorhab
315	33	55.9	461	7	ADG33762	Adg33762 Actinomyc	388	32	54.2	46	3	AB25736	Ab25736 Human sec
316	33	55.9	474	6	ABU11478	Abu11478 Human MDD	389	32	54.2	60	3	AAW59598	Aaw59598 Arabidops
317	33	55.9	488	6	ABU11731	Abu11731 Human MDD	390	32	54.2	60	3	AAG11176	Aag11176 Arabidops

391	32	54.2	65	2	AAR15046	Aar15046	C-termina	464	32	54.2	120	4	AAG64996	Aag64996	Nerve gro
392	32	54.2	66	8	ADG741135	Adg741135	BK polyom	465	32	54.2	120	4	AAG64999	Aag64999	Nerve gro
393	32	54.2	66	8	ABO55878	Abos5878	Human gen	466	32	54.2	120	4	AAG65011	Aag65011	Nerve gro
394	32	54.2	74	8	ADG741134	Adg741134	BK polyom	467	32	54.2	120	4	AAG65001	Aag65001	Nerve gro
395	32	54.2	78	4	AAW91088	Aaw91088	Human imm	468	32	54.2	120	4	AAG65013	Aag65013	Nerve gro
396	32	54.2	78	4	AAW91088	Aaw91088	Human imm	469	32	54.2	120	4	AAG65013	Aag65013	Nerve gro
397	32	54.2	78	4	AAW91088	Aaw91088	Human imm	470	32	54.2	120	4	AAG65013	Aag65013	Nerve gro
398	32	54.2	78	4	ABM49741	Abm49741	Propionib	471	32	54.2	120	5	ABG30438	Abg30438	Human ner
399	32	54.2	87	4	ABG16248	Abg16248	Nevel hum	472	32	54.2	120	5	ABG70692	Abg70692	Neuronal-
400	32	54.2	97	8	ADP21807	Adp21807	Human il6	473	32	54.2	120	6	ABM40896	Abm40896	Propionib
401	32	54.2	98	3	AGG00158	Aag00158	Human sec	474	32	54.2	120	6	ABP71158	Abp71158	Mouse mat
402	32	54.2	100	7	ADI40536	Adi40536	Human pur	475	32	54.2	120	6	ABP71156	Abp71156	Human mat
403	32	54.2	101	8	ADP21811	Adp21811	Human il6	476	32	54.2	120	6	ABP71157	Abp71157	Rat matur
404	32	54.2	107	7	ADC95446	Adc95446	E. faeciu	477	32	54.2	120	8	ADO99512	Ado99512	Human ner
405	32	54.2	111	3	AAG59597	Aag59597	Arabidops	478	32	54.2	121	2	AAR21872	Aar21872	Chimeric
406	32	54.2	112	4	AAG64420	Aag64420	Human pol	479	32	54.2	122	2	AAR21865	Aar21865	Chimeric
407	32	54.2	112	4	AAG64420	Aag64420	Human pol	480	32	54.2	122	2	AAR21852	Aar21852	Chimeric
408	32	54.2	115	5	ABJ04904	Abj04904	Human pro	481	32	54.2	123	2	AAR21860	Aar21860	Chimeric
409	32	54.2	115	4	ABG15671	Abg15671	Novel hum	482	32	54.2	124	2	AAR21851	Aar21851	Chimeric
410	32	54.2	117	6	ABR41549	Abra41549	Human dit	483	32	54.2	124	2	AAR21853	Aar21853	Chimeric
411	32	54.2	118	1	AAP91034	Aap91034	Human ner	484	32	54.2	124	2	AAR21854	Aar21854	Chimeric
412	32	54.2	118	2	AAR29493	Aar29493	NGF, mous	485	32	54.2	125	7	ABO64742	Abos64742	Klebeie11
413	32	54.2	118	2	AAR21874	Aar21874	Chimeric	486	32	54.2	129	2	AAR37539	Aar37539	Recombina
414	32	54.2	119	1	AAF40040	Aap40040	Sequence	487	32	54.2	129	2	AAW24145	Aaw24145	Recombina
415	32	54.2	119	1	AAP90133	Aap90133	Human ner	488	32	54.2	132	2	AAR21862	Aar21862	Chimeric
416	32	54.2	119	2	AAR77420	Aar77420	Human NGF	489	32	54.2	135	6	ABM70008	Abm70008	Photorhab
417	32	54.2	119	2	AAW04847	Aay04847	Mycobacte	490	32	54.2	152	5	AAW50302	Aam50302	Factor XI
418	32	54.2	119	3	AAW03347	Aab03347	Human bet	491	32	54.2	152	8	ADJ48934	Adj48934	Oil-assoc
419	32	54.2	120	2	AAR21866	Aar21866	Chimeric	492	32	54.2	153	4	AAG67676	Aag67676	Amino aci
420	32	54.2	120	2	AAR21873	Aar21873	Chimeric	493	32	54.2	153	6	ABG73918	Abg73918	Human ner
421	32	54.2	120	2	AAR21864	Aar21864	Chimeric	494	32	54.2	154	2	AAW22751	Aar22751	Human gro
422	32	54.2	120	2	AAR21867	Aar21867	Chimeric	495	32	54.2	156	5	AAW50303	Aam50303	Nerve gro
423	32	54.2	120	2	AAR21871	Aar21871	Chimeric	496	32	54.2	157	3	AAW01596	Aab01596	Nerve gro
424	32	54.2	120	2	AAR21868	Aar21868	Chimeric	497	32	54.2	157	4	AAG67677	Aag67677	Amino aci
425	32	54.2	120	2	AAR21863	Aar21863	Chimeric	498	32	54.2	157	5	AAW85725	Aaw85725	Synthetic
426	32	54.2	120	2	AAR21869	Aar21869	Chimeric	499	32	54.2	157	6	ABG73919	Abg73919	Human ner
427	32	54.2	120	2	AAW54084	Aar54084	Nerve gro	500	32	54.2	161	8	ADQ18900	Adq18900	Human sof
428	32	54.2	120	2	AAW90532	Aar90532	Pantropic				162	2	AAW41575	Aaw41575	Human car
429	32	54.2	120	2	AAW90531	Aar90531	Pantropic								
430	32	54.2	120	2	AAW48887	Aaw48887	Mouse ner								
431	32	54.2	120	2	AAW81125	Aaw81125	Nerve gro								
432	32	54.2	120	2	AAW81130	Aaw81130	Nerve gro								
433	32	54.2	120	2	AAW81133	Aaw81133	Nerve gro								
434	32	54.2	120	2	AAW81132	Aaw81132	Nerve gro								
435	32	54.2	120	2	AAW81126	Aaw81126	Nerve gro								
436	32	54.2	120	2	AAW81134	Aaw81134	Nerve gro								
437	32	54.2	120	2	AAW81124	Aaw81124	Nerve gro								
438	32	54.2	120	2	AAW81131	Aaw81131	Nerve gro								
439	32	54.2	120	2	AAW81127	Aaw81127	Nerve gro								
440	32	54.2	120	2	AAW81128	Aaw81128	Nerve gro								
441	32	54.2	120	2	AAW81121	Aaw81121	Nerve gro								
442	32	54.2	120	2	AAW81129	Aaw81129	Nerve gro								
443	32	54.2	120	2	AAW81122	Aaw81122	Nerve gro								
444	32	54.2	120	2	AAW81119	Aaw81119	Nerve gro								
445	32	54.2	120	2	AAW81120	Aaw81120	Nerve gro								
446	32	54.2	120	2	AAW81117	Aaw81117	Nerve gro								
447	32	54.2	120	2	AAW81117	Aaw81117	Nerve gro								
448	32	54.2	120	3	AAW29141	Aab29141	N-termina								
449	32	54.2	120	3	AAW22006	Aay22006	Human ner								
450	32	54.2	120	3	AAG59596	Aag59596	Arabidops								
451	32	54.2	120	3	AAU44177	Aau44177	Propionib								
452	32	54.2	120	4	AAG65003	Aag65003	Nerve gro								
453	32	54.2	120	4	AAG65002	Aag65002	Nerve gro								
454	32	54.2	120	4	AAG65005	Aag65005	Nerve gro								
455	32	54.2	120	4	AAG64994	Aag64994	Nerve gro								
456	32	54.2	120	4	AAG65004	Aag65004	Nerve gro								
457	32	54.2	120	4	AAG65006	Aag65006	Nerve gro								
458	32	54.2	120	4	AAG64997	Aag64997	Nerve gro								
459	32	54.2	120	4	AAG65000	Aag65000	Nerve gro								
460	32	54.2	120	4	AAG65014	Aag65014	Nerve gro								
461	32	54.2	120	4	AAG65010	Aag65010	Nerve gro								
462	32	54.2	120	4	AAG65012	Aag65012	Nerve gro								
463	32	54.2	120	4	AAG64998	Aag64998	Nerve gro								

ALIGNMENTS

RESULT 1

AAW79286	
ID	AAW79286 standard; protein; 311 AA.
AC	AAW79286;
XX	
XX	
DT	15-FEB-1999 (first entry)
DE	Human phosphatidic acid phosphatase beta.
DE	Human phosphatidic acid phosphatase beta.
XX	
KW	Phosphatidic acid phosphatase beta; PAP-beta; human; dephosphorylation; tumour suppressor; cancer; gene therapy.
OS	Homo sapiens.
PN	WO9846730-A1.
XX	
PD	22-OCT-1998.
XX	
PF	16-APR-1998; 98WO-US007928.
XX	
PR	17-APR-1997; 97US-00842827.
XX	
PA	(CELL-) CELL THERAPEUTICS INC.
PI	Leung DW, Tompkins CK;
XX	
DR	WPI: 1998-594568/50.
DR	N-PSDB; AAV69088.
PT	New nucleic acid encoding human phosphatidic acid phosphatases - used to

PT regulate levels of lipid cellular mediators and in gene therapy of e.g.
PT cancer.
XX Claim 8; Fig 3A-B; 62pp; English.
CC This is the amino acid sequence of human mature phosphatidic acid
CC phosphatase-beta (PAP-beta), an enzyme that catalyses the conversion of
CC phosphatidic acid into diacylglycerol. 3 Variants of human PAP, i.e. PAP-
CC alpha 1 and 2 (see AAW79284-85), PAP-beta (see AAW79286) and PAP-gamma
CC (see AAW79287) have been identified. The invention provides PAP
CC polynucleotides (see AAV69086-89) and polypeptides, a method of preparing
CC PAP in a transformed host cell, and a method of using PAP to
CC dephosphorylate a substrate, especially lysophosphatidic acid, ceramide 1
CC -phosphate or sphingosine 1-phosphate, particularly for production of
CC diacylglycerol, but also monoacylglycerol, ceramide and sphingosine. PAP
CC is able to control the balance of lipid mediators of cellular activation
CC and signal transduction. Sequences that encode PAP are potentially
CC useful, in gene therapy, for treatment of cancer (PAP may be tumour
CC suppressors, PAP-alpha is expressed at lower levels in cancer cells than
CC in normal cells of same tissue type), inflammatory disease and diabetes-
CC associated obesity
XX
SQ Sequence 311 AA;
Query Match 100.0%; Score 59; DB 2; Length 311;
Best Local Similarity 100.0%; Pred. No. 0.095;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NYRCRGDSDK 10
Db 178 NYRCRGDSDK 187
RESULT 2
ADBE0703
ID ADE60703 standard; protein; 311 AA.
XX ADE60703;
XX
XX 29-JAN-2004 (first entry)
XX Human Protein NP_003704, SEQ ID NO 6615.
XX Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX Homo sapiens.
XX WO2003016475-A2.
XX 27-FEB-2003.
XX 14-AUG-2002; 2002WO-US025765.
XX PF
XX 14-AUG-2001; 2001US-0312147P.
XX PR 01-NOV-2001; 2001US-0346382P.
XX PR 26-NOV-2001; 2001US-0333347P.
XX XX
XX (GEO) GEN HOSPITAL CORP.
XX (FARB) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
XX WPI; 2003-268312/26.
XX GENBANK; NP_003704.
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX Claim 1; Page; 1017pp; English.
XX The invention discloses a composition comprising two or more isolated rat

CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 311 AA;
Query Match 100.0%; Score 59; DB 7; Length 311;
Best Local Similarity 100.0%; Pred. No. 0.095;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NYRCRGDSDK 10
Db 178 NYRCRGDSDK 187
RESULT 3
ADM72103
ID ADM72103 standard; protein; 311 AA.
XX ADM72103;
XX
XX 03-JUN-2004 (first entry)
XX Human tau-related polypeptide PPAP2B.
XX
XX Tau; PPAP2A; PPAP2B; IHPK1; IHPK3; IHPK2; FLJ20530; DJ434014.5; EZF1;
KW Loc127424; Loc113179; KIAA0826; neuroprotective; nootropic;
KW antiparkinsonian; tau-protein kinase; human.
XX Homo sapiens.
XX WO2004022708-A2.
XX PN
XX 18-MAR-2004.
XX PD
XX 04-SEP-2003; 2003WO-US027590.
XX PF
XX 09-SEP-2002; 2002US-0408877P.
XX PR
XX (BGHM) BRIGHAM & WOMENS HOSPITAL INC.
XX PA
XX Feany MB, Shulman JM;
XX PI
XX WPI; 2004-248456/23.
XX DR
XX New substantially pure Tau-related polypeptides and polynucleotides,
PT useful for diagnosing and/or treating neurological conditions with
PT aberrant expression of the Tau-related polypeptide, such as Alzheimer's
PT disease.
XX

PS Claim 1; SEQ ID NO 2; 62pp; English.

XX The invention relates to tau-related polypeptide consisting essentially
 CC of an amino acid sequence selected from PPAP2A, PPAP2B, IHPK1, IHPK3,
 CC IHPK2, FLJ20530, D434014.5, EF1, LOC127424, LOC131179 and KIAA0826
 CC (ADMT7102-ADMT7112 respectively). The methods and compositions of the
 CC present invention are useful for the diagnosis and/or treatment of
 CC neurological diseases or conditions associated with aberrant expression
 CC or activity of the Tau-related polypeptide, such as Alzheimer's disease
 CC and Parkinson's disease. The present sequence represents a human tau-
 CC related polypeptide homologue PPAP2B.

XX Sequence 311 AA;

Query Match 100.0%; Score 59; DB 8; Length 311;

Best Local Similarity 100.0%; Pred. No. 0.095; 0; Indels 0; Gaps 0;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10

|||||

Db 178 NYRCRGDDSK 187

RESULT 4

ADN04415

ID ADN04415 standard; protein; 311 AA.

XX AC

ADN04415;

XX DT

01-JUL-2004 (first entry)

XX DE

Antipsoriatic protein sequence #401.

XX KW

antipsoriatic; gene therapy; psoriasis; diagnosis.

XX OS

Homo sapiens.

XX PN

WO2004028479-A2.

XX PD

08-APR-2004.

XX PF

25-SEP-2003; 2003WO-US030907.

XX PR

25-SEP-2002; 2002US-0414006P.

XX PA

(GETH) GENENTECH INC.

XX PI

Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;

PI Wu TD;

XX WPI;

2004-305105/28.

XX DR

N-PSDB; ADN04414.

XX PT

New PRO nucleic acid or polypeptide, useful for preparing a

PT pharmaceutical composition for diagnosing or treating psoriasis in a

PT mammal.

XX PS

Claim 9; SEQ ID NO 809; 3069pp; English.

XX CC

The invention relates to novel polynucleotide and polypeptides for
 CC treating psoriasis or a sequence having at least 80% identity to the
 CC above sequences. The nucleic acid is useful for preparing a composition
 CC for diagnosing or treating psoriasis in a mammal. This sequence
 CC corresponds to one of the polypeptides of the invention.

XX Sequence 311 AA;

Query Match

Best Local Similarity 100.0%; Score 59; DB 8; Length 311;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10

|||||

Db 178 NYRCRGDDSK 187

RESULT 5

ADP24274

ID ADP24274 standard; protein; 311 AA.

XX AC

ADP24274;

XX DT

18-NOV-2004 (first entry)

XX DE

PRO polypeptide SEQ ID NO:1452.

XX KW

PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;

osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;

antiaschmatic; hepatotropic; respiratory; gene therapy; immune system.

XX OS

Unidentified.

XX PN

WO2004041170-A2.

XX PD

21-MAY-2004.

XX PF

30-OCT-2003; 2003WO-US034312.

XX PR

01-NOV-2002; 2002US-0423394P.

XX PA

(GETH) GENENTECH INC.

XX PI

Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;

PI Wu TD;

XX WPI;

2004-419628/39.

XX DR

N-PSDB; ADP24273.

XX PT

New PRO polypeptides and polynucleotides, useful for treating e.g.
 PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated
 PT renal disease, or demyelinating diseases of the central or peripheral
 PT nervous system.

XX Claim 7; SEQ ID NO 1452; 2940pp; English.

XX The invention relates to a novel isolated nucleic acid and the PRO

polypeptide encoded by it. A protein of the invention has

antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,

osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,

antiaschmatic, hepatotropic, and respiratory activity. A polynucleotide

of the invention may have a use in gene therapy. The PRO polypeptide, its

agonist, antagonist, or antibody that specifically binds to the

polypeptide is useful for treating an immune related disorder such as

systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,

juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an

idiopathic inflammatory myopathy, Sjogren's syndrome, systemic

vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune

thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal

disease, a demyelinating disease of the central or peripheral nervous

system, idiopathic demyelinating polynuropathy, Guillain-Barre syndrome,

a chronic inflammatory demyelinating polynuropathy, a hepatobiliary

disease, infectious or autoimmune chronic active hepatitis, primary

biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,

inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's

disease, an autoimmune or immune-mediated skin disease, a bullous skin

disease, erythema multiforme, contact dermatitis, psoriasis, an allergic

disease, asthma, allergic rhinitis, atopic dermatitis, food

hypersensitivity, urticaria, an immunologic disease of the lung,

eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity

pneumonitis, a transplantation associated disease, graft rejection or

graft-versus-host disease. The present sequence represents a PRO protein

of the invention.

XX Sequence 311 AA;

Query Match 100.0%; Score 59; DB 8; Length 311;

Best Local Similarity 100.0%; Pred. No. 0.095;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10

Db 178 NYRCRGDDSK 187

RESULT 6

ABB99112
ID ABB99112 standard; protein; 312 AA.

XX

AC ABB99112;

XX 29-OCT-2002 (first entry)

DT

XX Rat phosphatase #2.

XX

XX Phosphatase; prenlyphosphoric acid; prenyl alcohol; rat.

XX

OS Rattus norvegicus.

XX

PN WO200253751-A1.

XX

PD 11-JUL-2002.

XX

PF 20-DEC-2001; 2001WO-JP011223.

XX

PR 28-DEC-2000; 2000JP-00401515.

XX

PR 28-DEC-2000; 2000JP-00401806.

XX

PA (TOYW) TOYOTA CHUO KENKYUSHO KK.

XX

XX (TOYT) TOYOTA JIDOSHA KK.

XX

PI Tokuhiko K, Muramoto N, Yamada Y, Asami O, Hirai M, Ohto C;

XX

PI Obata S, Muramatsu M;

XX

DR WPI; 2002-643303/69.

XX

DR N-PSDB; ABQ78948.

XX

XX Phosphatase polypeptides and encoded polynucleotides with substrate-

PT specificity for prenlyphosphoric acid, applicable in mass-production of

PT all-trans-prenyl alcohols by culturing transformants.

XX

PS Claim 1; Page 67-68; 93pp; Japanese.

XX

XX The invention relates to a novel phosphatase polypeptide, exhibiting

CC substrate-specificity for not less than 15C prenlyphosphoric acid. The

CC polynucleotides and encoded polypeptides are applicable in mass-

CC production of all-trans-prenyl alcohols. The sequence represents a rat

CC phosphatase polypeptide of the invention

XX

XX Sequence 312 AA;

QY

Query Match 93.2%; Score 55; DB 5; Length 312;

Best Local Similarity 90.0%; Pred. No. 0.45;

Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10

Db 179 NYRCRGDDSK 188

XX

AD660701 standard; protein; 312 AA.

XX

AC ADE60701;

XX

DT 29-JAN-2004 (first entry)

XX

XX Rat Protein CAA69106, SEQ ID NO 6613.

XX

KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XX Chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

OS Rattus norvegicus.

XX

PN WO2003016475-A2.

XX

PD 27-FEB-2003.

XX

PF 14-AUG-2002; 2002WO-US025765.

XX

PR 14-AUG-2001; 2001US-0312147P.

XX

PR 01-NOV-2001; 2001US-0346382P.

XX

XX 26-NOV-2001; 2001US-0333347P.

XX

PA (GEHO) GEN HOSPITAL CORP.

XX

PA (FARB) BAYER AG.

XX

PI Woolf C, D'urso D, Befort K, Costigan M;

XX

XX WPI; 2003-268312/26.

DR

DR GENBANK; CAA69106.

XX

PT New composition comprising two or more isolated polypeptides, useful for

PT preparing a medicament for treating pain in an animal.

XX

PS Claim 1; Page; 1017pp; English.

XX

XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 312 AA;

Query Match 93.2%; Score 55; DB 7; Length 312;

Best Local Similarity 90.0%; Pred. No. 0.45;

Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10

Db 179 NYRCRGDDSK 188

XX

RESULT 8

ABG23434

ID ABG23434 standard; protein; 152 AA.

XX

AC ABG23434;

XX

DT 18-FEB-2002 (first entry)
 XX Novel human diagnostic protein #23425.
 DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder.
 KW Homo sapiens.
 OS WO200175067-A2.
 PN 11-OCT-2001.
 PD 30-MAR-2001; 2001WO-US008631.
 XX 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX (HYSE-) HYSEQ INC.
 PA Drmanac RT, Liu C, Tang YT;
 PI WPI; 2001-639362/73.
 XX N-PSDB; AAS87621.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX Claim 20; SEQ ID NO 53793; 103pp; English.
 PS The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 152 AA;
 SQ Query Match 69.5%; Score 41; DB 4; Length 152;
 Best Local Similarity 70.0%; Pred. No. 50;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NYRCRGDDSK 10
 Db 68 NHRCRGDLTK 77
 |:||||| :|
 -:||||| :|
 RESULT 9
 AAU28253
 ID AAU28253 standard; protein; 237 AA.
 XX AAU28253;
 AC AAU28253;
 XX 18-DEC-2001 (first entry)
 XX
 DT

DE Novel human secretory protein, Seq ID No 610.
 XX Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
 KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
 KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
 KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
 KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
 KW gut protection; lung; liver fibrosis; immune deficiency; infection;
 KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
 KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
 KW fertility; analgesic; pain; antigen.
 XX Homo sapiens.
 OS WO200166689-A2.
 PN 13-SEP-2001.
 PD 05-MAR-2001; 2001WO-US004942.
 XX 07-MAR-2000; 2000US-00519705.
 PR 19-MAY-2000; 2000US-00574454.
 PR 17-JUN-2000; 2000US-00596193.
 PR 14-JUL-2000; 2000US-00616847.
 PR 19-SEP-2000; 2000US-00665363.
 PR 20-OCT-2000; 2000US-00693267.
 XX (HYSE-) HYSEQ INC.
 PA Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
 PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;
 XX WPI; 2001-589934/66.
 DR N-PSDB; AAS45153.
 XX Novel polypeptides and nucleic acids obtained from cDNA libraries
 PT prepared from various human tissues, for diagnosis and treatment of
 PT cancer, neurological, inflammatory, and autoimmune disorders.
 XX Example 2; SEQ ID NO 610; 107pp; English.
 PS The invention relates to novel isolated human secreted polypeptides (I)
 CC and polynucleotides (II). (I) and (II) are useful for treating
 CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
 CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
 CC involved in increasing haematopoiesis, stem cell survival, bone growth
 CC and remodeling. (I), (II) and modulators of (II) are useful for
 CC prophylaxis or treatment of one or more cancers. (II) is also useful for
 CC creating transgenic animals useful for studying the in vivo activities of
 CC the polypeptide as well as for studying modulators of the polypeptides.
 CC (I) induces the proliferation of neural cells and regeneration of nerve
 CC and brain tissue and is useful for the treatment of central and
 CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
 CC Parkinson's disease, Huntington's disease, and amyotrophic lateral
 CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
 CC activity, regulation of haematopoiesis and is useful for treating myeloid
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
 CC tissue growth, and in tissue repair, healing of burns, incisions, ulcers,
 CC for treating osteoporosis, osteoarthritis, bone degenerative disorders,
 CC or periodontal disease. Furthermore, (I) is also useful for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues, various immune deficiencies and
 CC disorders including severe combined immunodeficiency (SCID), bacterial or
 CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
 CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
 CC reactions and conditions, such as asthma or other respiratory problems.
 CC In addition, (I) affects biorhythms or circadian cycles of rhythms,
 CC fertility, metabolism, catabolism, anabolism, storage or elimination of
 CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
 CC analgesic effects or other pain reducing effects, immunoglobulin like
 CC activity and can act as an antigen in a vaccine composition to raise an
 CC immune response. AAU28020-AAU28395 represent novel human secreted protein

CC amino acid sequences of the invention

SQ Sequence 237 AA;

Query Match 69.5%; Score 41; DB 4; Length 237;

Best Local Similarity 70.0%; Pred. No. 78;

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10

||:|||||:|

Db 5 NHRCRGDLTK 14

RESULT 10

ADSL12199

ID ADSL12199 standard; protein; 237 AA.

XX AC

XX ADSL12199;

XX AC

DT 16-DEC-2004 (first entry)

XX XX

DE Human therapeutic contig protein - SEQ ID 2436.

XX KW

antiinflammatory; neuroprotective; antianaemic; cytostatic; vulnery;

KW inflammatory; haematopoiesis; immunity; neurodegenerative; stem cell;

KW aplastic anaemia; cancer; wound healing; gene therapy.

XX XX

OS Homo sapiens.

XX XX

FH Key Location/Qualifiers

FT Misc-difference 1..237

FT /label= Unknown, OTHER

FT /note= "OTHER = In-frame STOP codon"

XX XX

WO2004080148-A2.

XX XX

PD 23-SEP-2004.

XX XX

30-SEP-2003; 2003WO-US030720.

XX XX

02-OCT-2002; 2002US-0416186P.

XX XX

(NUVE-) NUVELO INC.

XX XX

PI Tang YT, Asundi V, Ren F, Zhang J, Wehrman T, Wang Z, Ma Y;

PI Wang D, Chen R, Zhao QA, Wang J, Ghosh M, Xue AJ, Weng G, Zhou P;

XX XX

WPI: 2004-668857/65.

DR N-PSDB; ADS11601.

XX XX

PT New polynucleotide, useful in preparing a composition for diagnosing or treating inflammatory, neurodegenerative or stem cell disorders, e.g., aplastic anemia or cancer for promoting wound healing.

XX XX

PS Example 2; SEQ ID NO 2436; 718pp; English.

XX XX

CC The invention relates to a novel isolated polynucleotide and the encoded polypeptide. The molecules of the invention demonstrate antiinflammatory, neuroprotective, antianaemic, cytostatic and vulnery activities and may be useful in preparing a composition for diagnosing or treating inflammatory, haematopoietic, immune, neurodegenerative or stem cell disorders, such as aplastic anaemia or cancer, as well as for promoting wound healing. The molecules may also be utilised during gene therapy procedures. The current sequence is that of a human therapeutic contig protein of the invention.

XX XX

SQ Sequence 237 AA;

Query Match 69.5%; Score 41; DB 8; Length 237;

Best Local Similarity 70.0%; Pred. No. 78;

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10

||:|||||:|

Db 5 NHRCRGDLTK 14

Db ||:|||||:|

5 NHRCRGDLTK 14

RESULT 11

ADJ50947

ID ADJ50947 standard; protein; 421 AA.

XX AC

XX ADJ50947;

XX XX

DT 06-MAY-2004 (first entry)

XX XX

DE Human novel protein NOV4b.

XX KW

Human; NOVX; autoimmune disease; Alzheimer's disease; stroke; allergy;

KW Parkinson's disease; Huntington's disease; multiple sclerosis; anxiety;

KW pain; diabetes; graft versus host disease; pancreatitis; obesity; ulcer;

KW anaemia; cancer; viral infection; bacterial infection;

KW parasitic infection.

XX XX

OS Homo sapiens.

XX XX

US2004030096-A1.

XX XX

PD 12-FEB-2004.

XX XX

01-AUG-2002; 2002US-00210281.

XX XX

02-AUG-2001; 2001US-0309501P.

PR 03-AUG-2001; 2001US-0310291P.

PR 08-AUG-2001; 2001US-0310951P.

PR 09-AUG-2001; 2001US-0311292P.

PR 13-AUG-2001; 2001US-0311979P.

PR 14-AUG-2001; 2001US-0312203P.

PR 17-AUG-2001; 2001US-0313201P.

PR 20-AUG-2001; 2001US-0313643P.

PR 20-AUG-2001; 2001US-0313702P.

PR 21-AUG-2001; 2001US-0314031P.

PR 23-AUG-2001; 2001US-0314466P.

PR 28-AUG-2001; 2001US-0315403P.

PR 29-AUG-2001; 2001US-0315853P.

PR 05-MAR-2002; 2002US-0361775P.

PR 05-MAR-2002; 2002US-0361832P.

XX XX

(GORM/) GORMAN L.

PA (ZERR/) ZERHUSEN B D.

PA (EDIN/) EDINGER S R.

PA (PADI/) PADIGARU M.

PA (GUOX/) GUO X.

PA (KEKU/) KEKUDA R.

PA (ZHON/) ZHONG M.

PA (PATT/) PATTURAJAN M.

PA (MILL/) MILLER C E.

PA (JIWU/) JI W.

PA (PENA/) PENA C E A.

PA (BURG/) BURGESS C E.

PA (SCIO/) SCIORE P.

PA (STON/) STONE D J.

PA (TAUP/) TAUPIER R J.

PA (CASH/) CASHMAN S J.

PA (ROTH/) ROTHENBERG M E.

PA (MAL/) MALYANKAR U M.

PA (BOLD/) BOLDOG F L.

XX XX

Gorman L, Zerhusen BD, Edinger SR, Padigaru M, Guo X, Kekuda R;

PI Zhong M, Patturajan M, Miller CE, Ji W, Pena, CEA, Burgess CE;

PI Sciore P, Stone DJ, Taupier RJ, Casman SJ, Rothenberg ME;

PI Malyankar UM, Boldog FL;

XX WPI: 2004-168942/16.

DR N-PSDB; ADJ50946.

XX XX

PT New NOVX polypeptides and polynucleotides, useful in diagnosing, treating

PT or preventing diseases or conditions, e.g. autoimmune disease,
PT Alzheimer's disease, diabetes, graft versus host disease, cancer or viral
PT or bacterial infections.
XX
PS Claim 2; SEQ ID NO 12; 342pp; English.
XX
CC The invention relates to an isolated NOVX polypeptide (of 44 disclosed)
CC comprising its mature form, a sequence having at least 95% sequence
CC identity to NOVX or a sequence comprising one or more conservative
CC substitutions in the amino acid sequence of NOVX. Also included are a
CC composition comprising NOVX and a carrier, a kit comprising, in one or
CC more containers, the composition, a method of identifying an agent that
CC binds to NOVX, a method for identifying a potential therapeutic agent for
CC use in treatment of a pathology related to aberrant expression or
CC aberrant physiological interactions of NOVX, a method for screening for a
CC modulator of activity of or of latency or predisposition to a pathology
CC associated with NOVX, a method for modulating the activity of NOVX, a
CC method of treating or preventing a pathology associated with NOVX or a
CC pathological state in a mammal, an isolated nucleic acid molecule
CC encoding a NOVX protein, a vector comprising the nucleic acid molecule,
CC a cell comprising the vector, an antibody that immunospecifically binds
CC to NOVX, a method for determining the presence or amount of NOVX or the
CC nucleic acid molecule in a sample, a method for determining the presence
CC of or predisposition to a disease associated with altered levels of
CC expression of NOVX or the nucleic acid molecule in a first mammalian
CC subject and a method of producing NOVX (comprising culturing the cell
CC under conditions that lead to expression of the polypeptide). NOVX is
CC useful in the manufacture of a medicament for treating a syndrome
CC associated with a human disease associated with NOVX. The polypeptides
CC and nucleic acid molecules are useful in diagnosing, treating or
CC preventing diseases or conditions, e.g. autoimmune disease, Alzheimer's
CC disease, stroke, allergies, Parkinson's disease, Huntington's disease,
CC multiple sclerosis, anxiety, pain, diabetes, graft versus host disease,
CC pancreatitis, obesity, ulcers, anaemia, cancer, viral or bacterial and
CC parasitic infections (many more diseases and disorders are listed in the
CC specification). The present sequence represents a NOVX protein.
XX
SQ Sequence 421 AA;
Query Match 69.5%; Score 41; DB 8; Length 421;
Best Local Similarity 70.0%; Pred. NO. 1.4e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 NVRCRGDSDK 10
Db 197 NHRCRGDLTK 206
:::|||||:
:::|||||:
RESULT 12
ADJ50945
ID ADJ50945 standard; protein; 424 AA.
XX
AC ADJ50945;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human novel protein NOV4a.
XX
KW Human; NOVX; autoimmune disease; Alzheimer's disease; stroke; allergy;
KW Parkinson's disease; Huntington's disease; multiple sclerosis; anxiety;
KW pain; diabetes; graft versus host disease; pancreatitis; obesity; ulcer;
KW anaemia; cancer; viral infection; bacterial infection;
KW parasitic infection.
XX
OS Homo sapiens.
XX
PN US2004030096-A1.
XX
PD 12-FEB-2004.
XX
PF 01-AUG-2002; 2002US-00210281.
XX
PR 02-AUG-2001; 2001US-0309501P.

PR 03-AUG-2001; 2001US-0310291P.
PR 08-AUG-2001; 2001US-0310951P.
PR 09-AUG-2001; 2001US-0311292P.
PR 13-AUG-2001; 2001US-0311973P.
PR 14-AUG-2001; 2001US-0312203P.
PR 17-AUG-2001; 2001US-0313201P.
PR 20-AUG-2001; 2001US-0313643P.
PR 20-AUG-2001; 2001US-0313702P.
PR 21-AUG-2001; 2001US-0314031P.
PR 23-AUG-2001; 2001US-0314466P.
PR 28-AUG-2001; 2001US-0315403P.
PR 29-AUG-2001; 2001US-0315853P.
PR 05-MAR-2002; 2002US-0361775P.
PR 05-MAR-2002; 2002US-0361832P.
XX
PA (GORM/) GORMAN L.
PA (ZERRH/) ZERRHUSEN B D.
PA (EDIN/) EDINGER S R.
PA (PADI/) PADIGARU M.
PA (GUOX/) GUO X.
PA (KEKU/) KEKUDA R.
PA (ZHON/) ZHONG M.
PA (PATI/) PATTURAJAN M.
PA (MILL/) MILLER C E.
PA (JIWW/) JI W.
PA (PENA/) PENA C E A.
PA (BURG/) BURGESS C E.
PA (SCIO/) SCIORE P.
PA (STON/) STONE D J.
PA (TAUP/) TAUPIER R J.
PA (CASM/) CASMAN S J.
PA (ROTH/) ROTHENBERG M E.
PA (MALY/) MALYANKAR U M.
PA (BOLD/) BOLDG F L.
XX
PI Gorman L, Zerrhusen BD, Edinger SR, Padigaru M, Guo X, Kekuda R;
PI Zhong M, Patturajan M, Miller CE, Ji W, Pena CEA, Burgess CE;
PI Sciore P, Stone DJ, Taupier RJ, Casman SJ, Rothenberg ME;
PI Malyankar UM, Boldog FL;
DR WPI; 2004-168942/16.
DR N-PSDB; ADJ50944.
XX
PT New NOVX polypeptides and polynucleotides, useful in diagnosing, treating
PT or preventing diseases or conditions, e.g. autoimmune disease,
PT Alzheimer's disease, diabetes, graft versus host disease, cancer or viral
PT or bacterial infections.
XX
PS Claim 2; SEQ ID NO 10; 342pp; English.
XX
CC The invention relates to an isolated NOVX polypeptide (of 44 disclosed)
CC comprising its mature form, a sequence having at least 95% sequence
CC identity to NOVX or a sequence comprising one or more conservative
CC substitutions in the amino acid sequence of NOVX. Also included are a
CC composition comprising NOVX and a carrier, a kit comprising, in one or
CC more containers, the composition, a method of identifying an agent that
CC binds to NOVX, a method for identifying a potential therapeutic agent for
CC use in treatment of a pathology related to aberrant expression or
CC aberrant physiological interactions of NOVX, a method for screening for a
CC modulator of activity of or of latency or predisposition to a pathology
CC associated with NOVX, a method for modulating the activity of NOVX, a
CC method of treating or preventing a pathology associated with NOVX or a
CC pathological state in a mammal, an isolated nucleic acid molecule
CC encoding a NOVX protein, a vector comprising the nucleic acid molecule,
CC a cell comprising the vector, an antibody that immunospecifically binds
CC to NOVX, a method for determining the presence or amount of NOVX or the
CC nucleic acid molecule in a sample, a method for determining the presence
CC of or predisposition to a disease associated with altered levels of
CC expression of NOVX or the nucleic acid molecule in a first mammalian
CC subject and a method of producing NOVX (comprising culturing the cell
CC under conditions that lead to expression of the polypeptide). NOVX is
CC useful in the manufacture of a medicament for treating a syndrome
CC associated with a human disease associated with NOVX. The polypeptides
CC and nucleic acid molecules are useful in diagnosing, treating or
CC preventing diseases or conditions, e.g. autoimmune disease, Alzheimer's
CC disease, stroke, allergies, Parkinson's disease, Huntington's disease,
CC multiple sclerosis, anxiety, pain, diabetes, graft versus host disease,
CC pancreatitis, obesity, ulcers, anaemia, cancer, viral or bacterial and
CC parasitic infections (many more diseases and disorders are listed in the
CC specification). The present sequence represents a NOVX protein.
XX

CC and nucleic acid molecules are useful in diagnosing, treating or
 CC preventing diseases or conditions, e.g. autoimmune disease, Alzheimer's
 CC disease, stroke, allergies, Parkinson's disease, Huntington's disease,
 CC multiple sclerosis, anxiety, pain, diabetes, graft versus host disease,
 CC pancreatitis, obesity, ulcers, anaemia, cancer, viral or bacterial and
 CC parasitic infections (many more diseases and disorders are listed in the
 CC specification). The present sequence represents a NOVX protein.

XX SQ Sequence 424 AA;
 Query Match 69.5%; Score 41; DB 8; Length 424;
 Best Local Similarity 70.0%; Pred. No. 1.4e+02;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NYRCRGDDSK 10
 : : : : : : :
 Db 198 NHRCRGDLTK 207

RESULT 13
 ADA83857
 ID ADA83857 standard; protein; 428 AA.
 XX AC
 XX AC ADA83857;
 XX DT 20-NOV-2003 (first entry)
 XX DE Human POM18 protein.
 XX KW human; marker; expressed sequence tag; EST; arabidopsis; tumour;
 KW stress-induced phenotype; hyperosmotic stress; colon cancer; immunogen;
 KW vaccine.
 XX KW
 XX OS Homo sapiens.
 XX WO2002103028-A2.
 XX PD 27-DEC-2002.
 XX PF 30-MAY-2002; 2002WO-IB004189.
 XX PR 30-MAY-2001; 2001US-0293999P.
 PR 22-OCT-2001; 2001US-0330457P.
 PR 19-FEB-2002; 2002US-0357144P.
 XX (BIOM-) BIOMEDICAL CENT.

XX Baranova AV, Yankovsky NK, Kozlov AP, Lobashev AV, Krukovskaya LL;
 PI WPI; 2003-175241/17.
 XX N-PSDB; ADA83856.
 XX Determining if a nucleic acid is a marker for a phenotype/cell type of
 PT interest, by global comparison of expressed sequence tags known to be
 PT expressed in the phenotype/cell type with all ESTs expressed in normal
 PT tissue.
 XX Claim 29; Page 269-271; 516pp; English.

XX The invention relates to a novel method for determining if a nucleic acid
 CC is a marker for a predetermined phenotype/cell type of interest from a
 CC biological species. The method comprises performing a global comparison
 CC of a group of expressed sequence tags (ESTs) known to be expressed in the
 CC phenotype/cell type of interest with all ESTs expressed in normal tissue
 CC in order to identify ESTs that are preferentially expressed in the
 CC phenotype/cell of interest. A method of the invention is useful for
 CC determining whether a nucleic acid is a marker for a predetermined
 CC phenotype or cell type of interest from a biological species, preferably
 CC Arabidopsis or human. The cell type of interest is an abnormal cell such
 CC as a tumour cell, and the predetermined phenotype is a stress-induced
 CC phenotype such as hyperosmotic stress or high salt conditions. A method
 CC of the invention is also useful for determining the progression of colon
 CC cancer in a human, for detecting a tumour cell, and for regulating or

CC preventing the growth of a tumour cell. An antibody of the invention is
 CC useful for detecting the absence or presence of peptides encoded by
 CC tumour-associated markers. A polypeptide of the invention is useful as an
 CC immunogen for vaccinating an animal. The present sequence represents a
 CC tumour-associated antigen of the invention.

XX SQ Sequence 428 AA;
 Query Match 69.5%; Score 41; DB 6; Length 428;
 Best Local Similarity 70.0%; Pred. No. 1.4e+02;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NYRCRGDDSK 10
 : : : : : : :
 Db 197 NHRCRGDLTK 206

RESULT 14
 ADR09234
 ID ADR09234 standard; protein; 481 AA.
 XX AC
 XX AC ADR09234;
 XX DT 04-NOV-2004 (first entry)
 XX DE Human protein useful for treating neurological disease Seq 2740.
 XX KW human; oligo-capping method; diagnostic marker; gene therapy;
 KW osteoporosis; neurological disease; Alzheimer's disease;
 KW Parkinson's disease; dementia; short memory; cancer;
 KW sense or motor function; emotional reaction; fear response; panic;
 KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;
 KW tranquiliser.
 XX OS Homo sapiens.
 XX EP1447413-A2.
 XX PD 18-AUG-2004.
 XX PF 12-FEB-2004; 2004EP-00003145.
 XX PR 14-FEB-2003; 2003JP-00102207.
 PR 09-MAY-2003; 2003JP-00131452.
 XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;
 PI Wakamatsu A, Ishii S, Nagai K, Irie R;
 XX WPI; 2004-583265/57.
 XX N-PSDB; ADR07278.
 XX New 1995 cDNA, useful for treating osteoporosis, neurological diseases,
 PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
 XX Claim 1; SEQ ID NO 2740; 2686pp; English.

XX This invention relates to novel, isolated full length human cDNA
 CC molecules and the encoded proteins thereof. Specifically, it refers to
 CC cDNA clones obtained by an oligo-capping method, where none of these
 CC clones are identical to any known human mRNAs. The present invention
 CC describes an immunoassay to identify agonists and antagonists, as well as
 CC antibodies, antisense molecules and siRNAs that can all be used to bind
 CC to and modulate expression of the cDNA molecules. As such, these
 CC molecules are useful for diagnostic markers or therapeutic targets for
 CC the various diseases or morbid states. In particular, they are useful in
 CC gene therapy for treating osteoporosis, neurological disease, Alzheimer's
 CC disease, Parkinson's disease, dementia, short memory and various cancers,
 CC as well as for maintaining equilibrium of sense or motor function, and
 CC for treating emotional reaction, fear response and panic. Accordingly,
 CC they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,
 CC cytostatic and tranquiliser activities. This polypeptide is a protein

CC encoded by a full length human cDNA sequence of the invention. NOTE: This
CC sequence is not given in the sequence listing of the specification but
CC can be obtained on CD-ROM from the European Patent Office, Vienna Sub-
CC office.
XX

XX SQ Sequence 481 AA;

Query Match 69.5%; Score 41; DB 8; Length 481;
Best Local Similarity 70.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
|:|||||:|
Db 250 NHRCRGDLTK 259

RESULT 15
ADS10966
ID ADS10966 standard; protein; 481 AA.

XX AC ADS10966;

XX DT 16-DEC-2004 (first entry)

XX DE Human therapeutic protein - SEQ ID 1203.

XX KW antiinflammatory; neuroprotective; antianaemic; cytostatic; vulnery;
XX KW inflammatory; haematopoesis; immunity; neurodegenerative; stem cell;
XX KW aplastic anaemia; cancer; wound healing; gene therapy.

XX OS Homo sapiens.

XX PN WO2004080148-A2.

XX PD 23-SEP-2004.

XX PF 30-SEP-2003; 2003WO-US030720.

XX PR 02-OCT-2002; 2002US-0416186P.

XX PA (NUVE-) NUVELO INC.

XX PI Tang YT, Asundi V, Ren F, Zhang J, Wehrman T, Wang Z, Ma Y;
XX PI Wang D, Chen R, Zhao QA, Wang J, Ghosh M, Xue AJ, Weng G, Zhou P;

XX WPI: 2004-668857/65.

XX DR N-PSDB; ADS10282.

XX PT New polynucleotide, useful in preparing a composition for diagnosing or
PT treating inflammatory, neurodegenerative or stem cell disorders, e.g.,
PT aplastic anemia or cancer for promoting wound healing.

XX PS Claim 20; SEQ ID NO 1203; 718pp; English.

XX CC The invention relates to a novel isolated polynucleotide and the encoded
CC polypeptide. The molecules of the invention demonstrate antiinflammatory,
CC neuroprotective, antianaemic, cytostatic and vulnery activities and may
CC be useful in preparing a composition for diagnosing or treating
CC inflammatory, haematopoietic, immune, neurodegenerative or stem cell
CC disorders, such as aplastic anaemia or cancer, as well as for promoting
CC wound healing. The molecules may also be utilised during gene therapy
CC procedures. The current sequence is that of a human therapeutic protein
CC of the invention. The current sequence is not shown explicitly within the
CC specification but can be accessed from the WIPO web-site.

XX SQ Sequence 481 AA;

Query Match 69.5%; Score 41; DB 8; Length 481;
Best Local Similarity 70.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
|:|||||:|

Db 250 NHRCRGDLTK 259

RESULT 16

ADM04303
ID ADM04303 standard; protein; 489 AA.

XX AC ADM04303;

XX DT 20-MAY-2004 (first entry)

XX DE Human protein of the invention SEQ ID NO:2988.

XX KW human; gene therapy; diagnostic marker; pharmaceutical.

XX OS Homo sapiens.

XX PN EP1347046-A1.

XX PD 24-SEP-2003.

XX PF 12-APR-2002; 2002EP-00008400.

XX PR 22-MAR-2002; 2002JP-00137785.

XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX PI Isogai T, Sugiyama T, Otsuki T, Wakanatsu A, Sato H, Ishii S;
XX PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
XX PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;

XX DR WPI: 2003-723558/69.

XX DR N-PSDB; ADM01860.

XX PT New polynucleotides and polypeptides are useful in gene therapy, for
PT developing a diagnostic marker or medicines for regulating their
PT expression and activity, or as a target of gene therapy.

XX PS Claim 1; SEQ ID NO 2988; 305pp; English.

XX CC The invention relates to a novel human polynucleotide and the encoded
CC polypeptide. A polynucleotide of the invention may have a use in gene
CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
CC as a primer for synthesizing the polynucleotide or as a probe for
CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are
CC useful in gene therapy, for developing a diagnostic marker or medicines
CC for regulating their expression and activity, or as a target of gene
CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
CC are useful as pharmaceutical agents. The present sequence represents a
CC protein sequence of the invention.

XX SQ Sequence 489 AA;

Query Match 69.5%; Score 41; DB 7; Length 489;
Best Local Similarity 70.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10
|:|||||:|

Db 258 NHRCRGDLTK 267

RESULT 17

ABU31582
ID ABU31582 standard; protein; 400 AA.

XX AC ABU31582;

XX DT 19-JUN-2003 (first entry)

XX DE Protein encoded by Prokaryotic essential gene #17109.

XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

```

XX OS Klebsiella pneumoniae.
XX AC W0200277183-A2.
XX DT 03-OCT-2002.
XX PD
XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX DR WPI; 2003-029926/02.
XX DR N-PSDB; ACA35452.
XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX PS Claim 25; SEQ ID NO 59506; 1766pp; English.
XX CC The invention relates to an isolated nucleic acid comprising any one of
XX CC the 6213 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX CC proliferation or the activity of a gene in an operon required for
XX CC proliferation; (7) identifying a compound that influences the activity of
XX CC the gene product or that has an activity against a biological pathway
XX CC required for proliferation, or that inhibits cellular proliferation; (8)
XX CC identifying a gene required for cellular proliferation or the biological
XX CC pathway in which a proliferation-required gene or its gene product lies
XX CC or a gene on which the test compound that inhibits proliferation of an
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX CC compound's activity; (11) a culture comprising strains in which the gene
XX CC product is overexpressed or underexpressed; (12) determining the extent
XX CC to which each of the strains is present in a culture or collection of
XX CC strains; or (13) identifying the target of a compound that inhibits the
XX CC proliferation of an organism. The antisense nucleic acids are useful for
XX CC identifying proteins or screening for homologous nucleic acids required
XX CC for cellular proliferation to isolate candidate molecules for rational
XX CC drug discovery programs, or for screening homologous nucleic acids
XX CC required for proliferation in cells other than S. aureus, S. typhimurium,
XX CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
XX CC the target prokaryotic essential genes. Note: The sequence data for this
XX CC patent did not form part of the printed specification, but was obtained
XX CC in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 400 AA;

Query Match 67.8%; Score 40; DB 6; Length 400;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDD 8
:|||||
Db 39 FRCRGDD 45

RESULT 18

```

```

ABU45262
ID ABU45262 standard; protein; 400 AA.
XX AC ABU45262;
XX DT 19-JUN-2003 (first entry)
XX DE Protein encoded by Prokaryotic essential gene #30789.
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX OS Salmonella paratyphi.
XX PN W0200277183-A2.
XX PD 03-OCT-2002.
XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX DR WPI; 2003-029926/02.
XX DR N-PSDB; ACA49132.
XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX PS Claim 25; SEQ ID NO 73186; 1766pp; English.
XX CC The invention relates to an isolated nucleic acid comprising any one of
XX CC the 6213 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX CC proliferation or the activity of a gene in an operon required for
XX CC proliferation; (7) identifying a compound that influences the activity of
XX CC the gene product or that has an activity against a biological pathway
XX CC required for proliferation, or that inhibits cellular proliferation; (8)
XX CC identifying a gene required for cellular proliferation or the biological
XX CC pathway in which a proliferation-required gene or its gene product lies
XX CC or a gene on which the test compound that inhibits proliferation of an
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX CC compound's activity; (11) a culture comprising strains in which the gene
XX CC product is overexpressed or underexpressed; (12) determining the extent
XX CC to which each of the strains is present in a culture or collection of
XX CC strains; or (13) identifying the target of a compound that inhibits the
XX CC proliferation of an organism. The antisense nucleic acids are useful for
XX CC identifying proteins or screening for homologous nucleic acids required
XX CC for cellular proliferation to isolate candidate molecules for rational
XX CC drug discovery programs, or for screening homologous nucleic acids
XX CC required for proliferation in cells other than S. aureus, S. typhimurium,
XX CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
XX CC the target prokaryotic essential genes. Note: The sequence data for this
XX CC patent did not form part of the printed specification, but was obtained
XX CC in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 400 AA;

```

Query Match 67.8%; Score 40; DB 6; Length 400;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDD 8
:|||||
DB 39 FRCRGDD 45

RESULT 19

ABU48051
ID ABU48051 standard; protein; 400 AA.

XX AC ABU48051;

XX DT 19-JUN-2003 (first entry)

XX DE Protein encoded by Prokaryotic essential gene #33578.

XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX OS Salmomella typhi.

XX PN WO200277183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-US009107.

XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.

XX PR 06-MAR-2002; 2002US-0362699P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX DR WPI; 2003-029926/02.

XX DR N-PSDB; ACA51921.

XX PS New antisense nucleic acids, useful for identifying proteins or screening

XX PT for homologous nucleic acids required for cellular proliferation to

XX PT isolate candidate molecules for rational drug discovery programs.

XX PS Claim 25; SEQ ID NO 75975; 1766pp; English.

XX CC The invention relates to an isolated nucleic acid comprising any one of
XX CC the 6213 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC polypeptide or its fragment whose expression is inhibited by the
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX CC proliferation or the activity of a gene in an operon required for
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX CC compound's activity; (11) a culture comprising strains in which the gene
XX CC product is overexpressed or underexpressed; (12) determining the extent
XX CC to which each of the strains is present in a culture or collection of
XX CC strains; or (13) identifying the target of a compound that inhibits the
XX CC proliferation of an organism. The antisense nucleic acids are useful for
XX CC identifying proteins or screening for homologous nucleic acids required
XX CC for cellular proliferation to isolate candidate molecules for rational

CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 400 AA;

Query Match 67.8%; Score 40; DB 6; Length 400;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDD 8

DB 39 FRCRGDD 45

RESULT 20

ABU15369

ID ABU15369 standard; protein; 400 AA.

XX AC ABU15369;

XX DT 19-JUN-2003 (first entry)

XX DE Protein encoded by Prokaryotic essential gene #896.

XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX OS Escherichia coli.

XX PN WO200277183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-US009107.

XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.

XX PR 06-MAR-2002; 2002US-0362699P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX DR WPI; 2003-029926/02.

XX DR N-PSDB; ACA19239.

XX PT New antisense nucleic acids, useful for identifying proteins or screening

XX PT for homologous nucleic acids required for cellular proliferation to

XX PT isolate candidate molecules for rational drug discovery programs.

XX PS Claim 25; SEQ ID NO 43293; 1766pp; English.

XX CC The invention relates to an isolated nucleic acid comprising any one of
XX CC the 6213 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC polypeptide or its fragment whose expression is inhibited by the
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX CC proliferation or the activity of a gene in an operon required for
XX CC proliferation; (7) identifying a compound that influences the activity of
XX CC the gene product or that has an activity against a biological pathway (8)
XX CC required for proliferation, or that inhibits cellular proliferation; (9)
XX CC identifying a gene required for cellular proliferation or the biological

CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 400 AA;

Query Match 67.8%; Score 40; DB 6; Length 400;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDD 8
DB 39 FRCRGDD 45
:|||||

RESULT 21
ABO65513
ID ABO65513 standard; protein; 424 AA.
AC ABO65513;
XX
XX 29-JUL-2004 (first entry)
XX
DE Klebsiella pneumoniae polypeptide seqid 12030.
XX
XX Recombinant expression vector; transcription regulatory element;
KW Klebsiella pneumoniae protein; antibacterial; Vaccine.
XX
XX Klebsiella pneumoniae.
OS
XX US6610836-B1.
XX
XX 26-AUG-2003.
XX
XX 27-JAN-2000; 2000US-00489039.
XX
XX 29-JAN-1999; 99US-0117747P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Breton GL, Osborne M;
XX
XX WPI; 2003-895346/82.
DR N-PSDB; ACH99064.
XX
XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
XX preparing a vaccine composition against Klebsiella pneumoniae.
XX
XX Disclosure; SEQ ID NO 12030; 932pp; English.
XX
XX The invention describes a new isolated nucleic acid encoding a Klebsiella
XX pneumoniae polypeptide. Also described are: a recombinant expression
XX vector comprising the nucleic acid, operably linked to a transcription
XX regulatory element; and a cell comprising the recombinant expression
XX vector. The nucleic acid is useful for preparing a vaccine composition
XX against Klebsiella pneumoniae. This is the amino acid sequence of a
XX Klebsiella pneumoniae polypeptide of the invention
XX
SQ Sequence 424 AA;

Query Match 67.8%; Score 40; DB 7; Length 424;
Best Local Similarity 85.7%; Pred. No. 2.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDD 8
DB 63 FRCRGDD 69
:|||||

RESULT 22
ADB74255
ID ADB74255 standard; protein; 705 AA.
XX
AC ADB74255;
XX
XX 04-DEC-2003 (first entry)
XX
XX Mycobacterium tuberculosis non-naturally occurring peptide #3.
XX
XX Non-naturally occurring peptide; anion pump protein; tuberculosis;
KW hypersensitivity reaction; tuberculostatic.
XX
XX Mycobacterium tuberculosis.
XX
XX US6583366-B1.
XX
XX 24-JUN-2003.
XX
XX 16-SEP-1994; 94US-00311731.
XX
XX 19-AUG-1993; 93US-00109181.
XX 22-OCT-1993; 93US-00142558.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Smith DR, Mao J;
XX
XX WPI; 2003-656441/62.
DR N-PSDB; ADB74252.
XX
XX New Mycobacterium tuberculosis anion pump peptide useful for as
XX tuberculosis vaccine and diagnosis of tuberculosis infection.
XX
XX Disclosure; SEQ ID NO 4; 26pp; English.
XX
XX The invention relates to a non-naturally occurring peptide of
XX Mycobacterium tuberculosis comprising an amino acid sequence
XX corresponding to an anion pump protein. The invention also relates to a
XX non-naturally occurring nucleic acid corresponding to a DNA sequence of
XX Mycobacterium tuberculosis or Mycobacterium leprae. The new peptide is
XX useful as a vaccine against Mycobacterium tuberculosis or Mycobacterium
XX leprae or for screening for new tuberculosis drugs. Purified proteins
XX derived from the sequences of the invention may elicit a specific immune
XX response. The peptide may also be used to detect hypersensitivity
XX reactions of individuals exposed to Mycobacterium tuberculosis or
XX Mycobacterium leprae. The proteins and peptides may be affixed to solid
XX supports to detect antibodies typical of hypersensitivity reactions, from
XX a patient's sera. This sequence represents Mycobacterium tuberculosis non
XX -naturally occurring peptide of the invention. Note: The sequence data
XX for this patent did not form part of the printed specification but was
XX obtained in electronic format directly from USPTO at
XX seqdata.uspto.gov/sequence.html.
XX
XX Sequence 705 AA;

Query Match 67.8%; Score 40; DB 7; Length 705;
Best Local Similarity 66.7%; Pred. No. 3.4e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 YRCRGDDSK 10
DB 383 YRCRADDT 391
:|||||

```
RESULT 23
ABP72189
ID ABP72189 standard; protein; 1304 AA.
XX
AC ABP72189;
XX
DT 22-APR-2003 (first entry)
XX
DE Plasmodium berghei modular secreted protein PbSR.
XX
KW PbSR; protein secretion; malaria; diagnosis; vaccine; antimalarial;
KW protozoacide.
XX
OS Plasmodium berghei.
XX
FH Key Location/Qualifiers
FT Peptide 1..22
FT /label= Signal_peptide
FT Protein 23..1304
FT /label= Mature_protein
XX
PN WO2003004524-A2.
XX
XX 16-JAN-2003.
XX
PF 02-JUL-2002; 2002WO-GB003045.
XX
XX 02-JUL-2001; 2001GB-00016185.
XX
XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
XX
PI Claudianos C, Crompton TK, Dessens JT, Sindén RE, Trueman HE;
XX
WPI; 2003-210339/20.
DR N-PSDB; ABZ58196.
XX
XX New secreted proteins from malarial parasites (e.g. Plasmodium
PT falciparum), useful for detecting or diagnosing malaria, or as a vaccine
PT or medicament for the treatment and/or prophylaxis of malaria.
XX
XX Claim 1; Fig 1; 67pp; English.
XX
CC The present sequence is the protein sequence of novel secreted protein
CC PbSR from the mouse malaria parasite, Plasmodium berghei. The invention
CC provides novel modular secreted proteins from Plasmodium falciparum
CC (PfSR), P. berghei (PbSR) and Plasmodium yoelii (PySR) that are post-
CC translationally modified in sporozoites. These include Limulus clotting
CC factor, lipid binding, scavenger receptor cysteine-rich and pentraxin
CC domains. Homologues were identified in Plasmodium knowlesi (PkSR),
CC Plasmodium chabaudi (PcSR), Cryptosporidium parvum (CpSR), Toxoplasma
CC gondii (TgSR) and Theileria parva (TpSR). PbSR, PfSR and PySR are from a
CC unique and previously unrecognized family containing domains associated
CC with binding and modulating host proteins involved in immunity. They are
CC the first proteins from Plasmodium which are secreted and which show a
CC gene disruption phenotype in both vertebrate and mosquito, making them
CC highly useful for the production of malaria vaccines. Transgenic PbSR
CC null mutants have attenuated growth in immunocompetent mice. They fail to
CC produce sporozoites in Anopheles stephensi mosquitoes and consequently
CC are dead-end parasites in the midgut of the mosquito. The novel
CC Plasmodium proteins, and the nucleic acids encoding them, are also in
CC claimed methods for the detection/diagnosis of malaria, and are used
CC in vaccine compositions. Agents capable of antagonising, inhibiting or
CC interfering with the function or expression of the secreted protein are
CC used in the treatment and/or prophylaxis of malaria
XX
XX Sequence 1304 AA;
XX
Query Match 67.8%; Score 40; DB 6; Length 1304;
Best Local Similarity 55.6%; Pred. No. 6.2e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
XX
XX
RESULT 24
ABP72191
ID ABP72191 standard; protein; 451 AA.
XX
AC ABP72191;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 30384.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
XX
XX 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
WPI; 2001-658860/75.
DR N-PSDB; ABL11967.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX Disclosure; SEQ ID NO 30384; 2lpp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 451 AA;
XX
Query Match 66.1%; Score 39; DB 4; Length 451;
Best Local Similarity 75.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
XX
XX
QY 2 YRCRGDD5 9
DB 3 YRCRGNES 10
XX
XX
RESULT 25
ABP72191
ID ABP72191 standard; protein; 1299 AA.
XX
AC ABP72191;
XX
DT 22-APR-2003 (first entry)
XX
XX Plasmodium yoelii modular secreted protein PySR.
```

XX PfSR; protein secretion; malaria; diagnosis; vaccine; antimalarial;
 KW protozoacide.
 XX Plasmodium yoelii.
 XX
 XX Key Location/Qualifiers
 FH Peptide 1..22 /label= Signal_peptide
 FT Protein 23..2399 /label= Mature_protein
 FT Misc-difference 175 /note= "encoded by ATK"
 FT
 XX WO2003004524-A2.
 XX
 XX 16-JAN-2003.
 XX
 XX 02-JUL-2002; 2002WO-GB003045.
 XX
 XX 02-JUL-2001; 2001GB-00016185.
 XX
 XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
 XX
 XX Claudianos C, Crompton TK, Dessens JT, Sindén RE, Trueman HE;
 XX
 XX WPI; 2003-210339/20.
 XX N-PSDB; ABZ58198.
 XX
 XX New secreted proteins from malarial parasites (e.g. Plasmodium
 PT falciparum), useful for detecting or diagnosing malaria, or as a vaccine
 PT or medicament for the treatment and/or prophylaxis of malaria.
 XX
 XX Claim 1; Fig 1; 67pp; English.
 XX
 XX The present sequence is the protein sequence of novel secreted protein
 CC PySR from the mouse malaria parasite, Plasmodium yoelii. The invention
 CC provides novel modular secreted proteins from Plasmodium falciparum
 CC (PfSR), Plasmodium berghei (PbSR) and P. yoelii (PySR) that are post-
 CC translationally modified in sporozoites. These include Limulus clotting
 CC factor, lipid binding, scavenger receptor cysteine-rich and pentraxin
 CC domains. Homologues were identified in Plasmodium knowlesi (PkSR),
 CC Plasmodium chabaudi (PcSR), Cryptosporidium parvum (CpSR), Toxoplasma
 CC gondii (TgSR) and Theileria parva (TpSR). PbSR, PfSR and PySR are from a
 CC unique and previously unrecognized family containing domains associated
 CC with binding and modulating host proteins involved in immunity. They are
 CC the first proteins from Plasmodium which are secreted and which show a
 CC gene disruption phenotype in both vertebrate and mosquito, making them
 CC highly useful for the production of malaria vaccines. Transgenic PbSR
 CC null mutants have attenuated growth in immunocompetent mice. They fail to
 CC produce sporozoites in Anopheles stephensi mosquitoes and consequently
 CC are dead-end parasites in the midgut of the mosquito. The novel
 CC Plasmodium proteins, and the nucleic acids encoding them, are used in
 CC claimed methods for the detection/diagnosis of malaria, and are also used
 CC in vaccine compositions. Agents capable of antagonising, inhibiting or
 CC interfering with the function or expression of the secreted protein are
 CC used in the treatment and/or prophylaxis of malaria
 XX
 XX Sequence 1299 AA;
 XX
 XX Query Match 66.1%; Score 39; DB 6; Length 1299;
 XX Best Local Similarity 62.5%; Pred. No. 9.2e+02;
 XX Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 1 NYRCRGDD 8
 XX |.:|:|:|:
 XX Db 614 NFRCKGDE 621
 XX
 XX RESULT 26
 XX AAW99087
 XX ID AAW99087 standard; protein; 822 AA.
 XX

AC AAW99087;
 XX
 XX 13-MAY-1999 (first entry)
 XX
 XX Human serine protease BSSP-3.
 XX
 XX Serine protease; BSSP-3; brain tissue.
 XX
 XX Homo sapiens.
 XX
 XX WO9905290-A1.
 XX
 XX 04-FEB-1999.
 XX
 XX 24-JUL-1998; 98WO-JP003324.
 XX
 XX 24-JUL-1997; 97JP-00213969.
 XX
 XX (SUNR) SUNTORY LTD.
 XX
 XX Tsuruoka N, Yamashiro K, Yamaguchi N;
 XX
 XX WPI; 1999-142942/12.
 XX N-PSDB; AAX19024.
 XX
 XX New serine protease expressed in brain tissue - used in screening for
 PT potential serine protease inhibitors for drug use.
 PT
 XX Claim 1; Page 61-65; 69pp; Japanese.
 XX
 XX The present sequence is a serine protease designated BSSP-3, which is
 CC isolated from human brain tissue. Transforms may be used to produce
 CC the enzyme or its partial sequences. Products from the present invention
 CC are used for screening for potential peptide or non-peptide serine
 CC protease inhibitors or expression regulators for use as drugs
 XX
 XX Sequence 822 AA;
 XX
 XX Query Match 64.4%; Score 38; DB 2; Length 822;
 XX Best Local Similarity 75.0%; Pred. No. 8.6e+02;
 XX Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 XX
 XX QY 1 NYRCRGDD 8
 XX |.:|:|:|:
 XX Db 183 NVRCKGDE 190
 XX
 XX RESULT 27
 XX AAW83361
 XX ID AAW83361 standard; protein; 875 AA.
 XX
 XX AAW83361;
 XX
 XX 17-FEB-1999 (first entry)
 XX
 XX Human neurotrypsin.
 XX
 XX Human; neurotrypsin; tumour inhibition; neurological disease;
 KW lung disease; gene therapy; drug development; stroke; brain injury;
 KW neurodegeneration; neuroinflammatory disease; multiple sclerosis;
 KW epilepsy; hypoxia; ischaemia; nerve transection; neovascularisation;
 KW emphysema; bronchitis.
 XX
 XX Homo sapiens.
 XX
 XX WO9849322-A1.
 XX
 XX 05-NOV-1998.
 XX
 XX 24-APR-1998; 98WO-IB000625.
 XX
 XX 26-APR-1997; 97CH-00000966.
 XX

PA	(SOND/) SONDEREGGER P.
XX	Sonderegger P;
PI	
XX	
DR	WPI; 1999-009438/01.
DR	N-PSDB; AAV72589.
XX	
PT	New human and murine neurotrophin - used, e.g. for inhibiting tumours,
PT	treatment of neurological or lung disease, including by gene therapy and
PT	in drug development.
XX	
PS	Claim 1; Page 20-24; 50pp; English.
XX	
CC	The present sequence represents human neurotrophin. Neurotrophin proteins
CC	and polynucleotides can be used: (i) to inhibit tumours, including
CC	metastases, e.g. of brain or retina; (ii) to minimise tissue damage
CC	caused by stroke or brain injury (having a protective effect on the
CC	pneumonia zone); (iii) to treat or prevent neurodegeneration,
CC	neuroinflammatory disease (e.g. multiple sclerosis) or epilepsy; (iv) to
CC	increase survival of damaged neurons (e.g. in cases of hypoxia,
CC	ischaemia, nerve transection) and to stimulate regeneration and/or
CC	restoration of synapses; (v) to treat or prevent retinal disorders (e.g.
CC	degeneration or neovascularisation); (vi) to prevent apoptosis (or other
CC	causes of cell death) in the nervous system; (vii) to regenerate brain
CC	and/or nervous tissue; (viii) to treat pain; (ix) to improve brain
CC	performance, including learning and memory; (x) to treat or prevent a
CC	wide range of psychiatric disorders; and (xi) to treat brain or lung
CC	injury associated with protease expression (specifically emphysema or
CC	bronchitis)
XX	
XX	
SQ	Sequence 875 AA;
	Query Match 64.4%; Score 38; DB 2; Length 875;
	Best Local Similarity 75.0%; Pred No. 9.1e+02;
	Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0
Qy	1 NVRCRGDD 8 236 NVRCRGDE 243
Dd	
RESULT 28	
ADQ21226	
ID	ADQ21226 standard; protein; 875 AA.
AC	ADQ21226;
XX	
DT	26-AUG-2004 (first entry)
XX	
DE	Human soft tissue sarcoma-upregulated protein - SEQ ID 4046.
DE	
KM	soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.
OS	Homo sapiens.
XX	
PN	WO2004048938-A2.
PD	
PD	10-JUN-2004.
XX	
PF	26-NOV-2003; 2003WO-US038193.
XX	
PR	26-NOV-2002; 2002US-0429739P.
XX	
PA	(PROT-) PROTEIN DESIGN LABS INC.
XX	
PI	Aziz N, Ginsburg WM, Zlotnik A;
XX	
DR	WPI; 2004-441208/41.
XX	
PT	Early detection of soft tissue sarcoma comprises determining expression
PT	of a gene in a first soft tissue sample and a normal soft tissue sample
PT	and comparing the gene expression, also useful in treating soft tissue
PT	sarcoma.

XX The present invention describes the use of polypeptides related to
 CC urological disorders for identifying a compound capable of treating a
 CC urological disorder, identifying a subject having a urological disorder,
 CC or treating a subject having a urological disorder. Also described: (1) a
 CC method for identifying a compound capable of treating a urological
 CC disorder; (2) a method for identifying a subject having a urological
 CC disorder; and (3) a method for treating a subject having a urological
 CC disorder. The compound has urothelial and cytostatic activities. The
 CC polypeptides related to urological disorders are useful for identifying a
 CC compound capable of treating a urological disorder, identifying a subject
 CC having a urological disorder, or treating a subject having a urological
 CC disorder. Disorders include urinary incontinence and benign prostatic
 CC hyperplasia. The present sequence represents a human urological disorder
 CC related protein, which is used in the exemplification of the present
 CC invention.
 XX
 SQ Sequence 875 AA;
 Query Match 64.4%; Score 38; DB 8; Length 875;
 Best Local Similarity 75.0%; Pred. No. 9.1e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NVRCRGDD 8
 Db 236 NVRCRGDE 243
 RESULT 30
 ABP42738
 ID ABP42738 standard; protein; 77 AA.
 XX
 AC ABP42738;
 XX
 DT 22-AUG-2002 (first entry)
 XX
 DE Human ovarian antigen HOVDU91, SEQ ID NO:3870.
 XX
 KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW antiinflammatory; gynaecological; reproductive.
 XX
 OS Homo sapiens.
 XX
 PN WO200200677-A1.
 XX
 PD 03-JAN-2002.
 XX
 PF 07-JUN-2001; 2001WO-US018569.
 XX
 PR 07-JUN-2000; 2000US-0209467P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Birse CE, Rosen CA;
 XX
 DR WPI; 2002-147878/19.
 DR N-PSDB; ABQ55815.
 XX
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
 PT cancer), immune disorders, cardiovascular disorders and neurological
 PT diseases.
 XX
 PS Claim 11; SEQ ID NO 3870; 2922pp; English.
 XX

CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), and toxic
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents a human ovarian antigen of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 77 AA;
 Query Match 62.7%; Score 37; DB 5; Length 77;
 Best Local Similarity 60.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 NVRCRGDDSK 10
 Db 41 NFFCRGPDGSK 50
 RESULT 31
 AAU61299
 ID AAU61299 standard; protein; 92 AA.
 XX
 AC AAU61299;
 XX
 DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #22195.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; EUSA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US012865.
 XX
 PR 21-APR-2000; 2000US-0199047P.
 PR 02-JUN-2000; 2000US-0208841P.
 PR 07-JUL-2000; 2000US-0216747P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonmeuve J, Zhang Y, Jen S, Carter D;
 XX

DR WPI; 2001-616774/71.
 DR N-PSDB; AAS59616.
 XX
 PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.
 XX
 XX Example 1; SEQ ID NO 22494; 1069pp; English.
 XX
 XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 92 AA;
 Query Match 62.7%; Score 37; DB 4; Length 92;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 RCRGDD 8
 Db 74 RCRGDD 79
 RESULT 32
 ABM57818
 ID ABM57818 standard; protein; 92 AA.
 AC ABM57818;
 XX
 XX 20-OCT-2003 (first entry)
 DT
 DE Propionibacterium acnes predicted ORF-encoded polypeptide #22494.
 XX
 XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;
 KW immunostimulant; immune response; vaccine.
 KW
 XX Propionibacterium acnes.
 OS
 XX
 XX WO2003033515-A1.
 PN
 XX 24-APR-2003.
 PD
 XX 11-OCT-2002; 2002WO-US032727.
 PF
 XX 15-OCT-2001; 2001US-00978825.
 PR
 XX (CORI-) CORIXA CORP.
 PA
 XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
 PI Zhang Y, Wang S, Jen S, Lodes WJ, Benson DR, Jones R, Carter D;
 PI Barth B, Vallieue-Douglass J;
 XX
 XX WPI; 2003-381789/36.
 DR N-PSDB; ACF64545.
 XX

PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
 PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
 PT or for stimulating an immune response specific for a P. acnes protein.
 XX
 XX Example 1; SEQ ID NO 22494; 1481pp; English.
 XX
 XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
 CC encoding a Propionibacterium acnes protein. The invention also relates to
 CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
 CC immunogenic fragments of P. acnes polypeptides. The invention
 CC additionally encompasses expression vectors and host cells comprising a
 CC polynucleotide of the invention; antibodies against polypeptides of the
 CC invention; fusion proteins comprising a polypeptide of the invention; a
 CC method for stimulating an immune response specific for a P. acnes
 CC polypeptide and an isolated T cell population comprising P. acnes polypeptides,
 CC via this method; a vaccine composition (comprising P. acnes polypeptides,
 CC polynucleotides, antibodies, fusion proteins, T cell populations, or
 CC antigen-presenting cells that express the polypeptide); a method and kit
 CC for detecting or determining the presence or absence of P. acnes in a
 CC patient; and a method for inhibiting the development of P. acnes in a
 CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
 CC proteins, T cell populations or antigen-presenting cells that express the
 CC polypeptides are useful for diagnosing, preventing or treating acne
 CC vulgaris, or for stimulating an immune response specific for a P. acnes
 CC protein. The polynucleotides can also be used as probes or primers for
 CC nucleic acid hybridisation. The vaccine composition is useful for the
 CC stimulation of an immune response against P. acnes, or for treating acne,
 CC and the kit is useful for performing a diagnostic assay. The present
 CC sequence represents a polypeptide predicted to be encoded by an ORF (open
 CC reading frame) contained within the P. acnes polynucleotides of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 92 AA;
 Query Match 62.7%; Score 37; DB 6; Length 92;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 RCRGDD 8
 Db 74 RCRGDD 79
 RESULT 33
 AAM91935
 ID AAM91935 standard; protein; 113 AA.
 AC AAM91935;
 XX
 XX 05-NOV-2001 (first entry)
 DT
 XX Human digestive system antigen SEQ ID NO: 1284.
 DE
 XX Human; digestive system antigen; gene therapy; cancer; appendicitis;
 KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
 KW digestive system disorder; Meckel's diverticulum.
 XX
 XX Homo sapiens.
 OS
 XX WO200155314-A2.
 PN
 XX 02-AUG-2001.
 PD
 XX 17-JAN-2001; 2001WO-US001324.
 PF
 XX 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR

CC diagnosis, treatment and prevention of digestive system disorders,
CC including cancer, Meckel's diverticulum, bacterial or parasitic
CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
CC ulcerative colitis. The present sequence is a digestive system antigen of
CC the invention
XX
SQ Sequence 113 AA;

Query Match 62.7%; Score 37; DB 4; Length 113;
Best Local Similarity 62.5%; Pred. No. 1.8e+02;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NYRCRGDD 8
||:|:|:
Db 10 NVQCEGED 17

RESULT 34

AAM38569
ID AAM38569 standard; protein; 113 AA.

XX AC AAM38569;

XX DT 19-OCT-2001 (first entry)

XX DE Human colorectal cancer antigen SEQ ID NO: 84.

XX KW Human; colorectal cancer; colorectal cancer antigen; gene therapy.

XX OS Homo sapiens.

XX PN WO200155350-A1.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001350.

XX PR 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 14-JUL-2000; 2000US-0217496P.

PR 26-JUL-2000; 2000US-0218293P.

PR 26-JUL-2000; 2000US-0220963P.

PR 14-AUG-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226277P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226686P.

PR 23-AUG-2000; 2000US-0227182P.

PR 30-AUG-2000; 2000US-0227009P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246603P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.


```

PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249219P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX WPI; 2001-457727/49.
DR N-PSDB; AA157547.
XX
XX Isolated polypeptide for treating, preventing and/ or prognosing
PT disorders related to the colon and rectum including colorectal cancers
PT and also for testing and detection e.g. diagnosis.
XX
XX Claim 11; SEQ ID NO 84; 522pp + Sequence Listing; English.
XX
XX The present invention provides the protein and coding sequences of a
CC number of colorectal cancer antigens. These are shown in AA157547-
CC AA157619 and AA157619-AMW38641. These can be used in the diagnosis,
CC prevention and treatment of cancer of the colon and/or rectum. The
CC present sequence is a colorectal cancer antigen of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 113 AA;
Query Match 62.7%; Score 37; DB 4; Length 113;
Best Local Similarity 62.5%; Pred. No. 1.8e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 NYRCRGDD 8
DB 10 NYQCEGED 17
XX
XX RESULT 35
ABG97621
ID ABG97621 standard; protein; 113 AA.
XX
XX AC ABG97621;
XX
XX DT 18-DEC-2002 (first entry)
XX
XX DE Human colorectal cancer related protein #1.
XX
XX Human; colorectal cancer related protein; colon; rectum;
KW colorectal cancer metastasis; gastrointestinal disorder; cytostatic.

```

```

XX Homo sapiens.
XX US2002119919-A1.
XX 29-AUG-2002.
XX
XX 17-JAN-2001; 2001US-00764855.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX
XX (ROSE/) ROSEN C A.
XX (RUBE/) RUBEN S M.
XX (BARA/) BARASH S C.
XX
XX Rosen CA, Ruben SM, Barash SC;
PI
XX WPI; 2002-731367/79.
DR N-PSDB; ABS99724.
XX
XX New colorectal cancer polypeptide for diagnosing, prognosing, preventing,
PT and treating immune, hyperproliferative, liver, kidney, reproductive
PT disorders and for identifying modulators of therapeutic use.
XX
XX Claim 11; SEQ ID NO 84; 183pp; English.
XX
XX The present invention relates to the isolation of novel human colorectal
CC cancer related proteins, and polynucleotide sequences encoding them. The
CC sequences of the invention are useful in the diagnosis, treatment,
CC prevention and/or prognosis of the colon and/or rectum, including
CC colorectal cancer, colorectal cancer metastases, and gastrointestinal
CC disorders such as dysphagia, peptic oesophagitis, gastric reflux,
CC irritable bowel syndrome, and peritoneal diseases. The invention also
CC describes antibodies that bind colorectal cancer related proteins,
CC vectors, host cells, and recombinant and synthetic methods for producing
CC human colorectal cancer related polynucleotides, polypeptides, and/or
CC antibodies. ABG97621-ABG97693 represent human colorectal cancer related
CC proteins. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from the USPTO web site at seqdata.uspto.gov/psipsIDEntry.html
XX
XX SQ Sequence 113 AA;
Query Match 62.7%; Score 37; DB 5; Length 113;
Best Local Similarity 62.5%; Pred. No. 1.8e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 NYRCRGDD 8
DB 10 NYQCEGED 17
XX
XX RESULT 36
ADB92877
ID ADB92877 standard; protein; 113 AA.
XX
XX AC ADB92877;
XX
XX DT 04-DEC-2003 (first entry)
XX
XX DE Human colorectal cancer related polypeptide #1.
XX
XX human; colorectal cancer; antigen; gene therapy;
KW gastrointestinal disorder; inflammatory disease; infection; cancer;
KW intestinal neoplasm; small intestine carcinoma; tumour;
KW small intestine non-Hodgkin's lymphoma; small bowel lymphoma; ulcer;
KW peptic ulcer; Bruton's disease; X linked infantile agammaglobulinaemia;
KW severe combined immunodeficiency; DiGeorge anomaly;
KW hyperproliferative disorder; acute lymphoblastic leukaemia;
KW acute lymphocytic leukaemia; urinary system disorder; cortical necrosis;
KW kidney infarction; cardiovascular disorder; carcinoid heart disease;
KW arrhythmia; respiratory disorder; non-allergic rhinitis; sinusitis;
KW musculoskeletal system disorder; Albers-Schonberg disease;

```

KW Marfan's syndrome; neurological disease; phenylketonuria;
KW Wernicke's encephalopathy; Alzheimer's disease; endocrine disorder;
KW Grave's disease; Cushing's syndrome; reproductive system disorder;
KW prostatitis; benign prostatic hypertrophy; benign prostatic hyperplasia;
KW thrombosis; atherosclerosis; myocardial infarction; ischaemic attack.

XX Homo sapiens.

OS US2003054420-A1.

PN 20-MAR-2003.

PD 11-FEB-2002; 2002US-00072349.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226868P.

PR 22-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227009P.

KW Marfan's syndrome; neurological disease; phenylketonuria;
KW Wernicke's encephalopathy; Alzheimer's disease; endocrine disorder;
KW Grave's disease; Cushing's syndrome; reproductive system disorder;
KW prostatitis; benign prostatic hypertrophy; benign prostatic hyperplasia;
KW thrombosis; atherosclerosis; myocardial infarction; ischaemic attack.

XX Homo sapiens.

OS US2003054420-A1.

PN 20-MAR-2003.

PD 11-FEB-2002; 2002US-00072349.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226868P.

PR 22-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227009P.

PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-023935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.

PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 PR 17-JAN-2001; 2001US-00764855.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM, Barash SC;
 XX
 DR WPI; 2003-708345/67.
 DR N-PSDB; ADB92794.
 XX
 PT Novel colorectal cancer antigen useful for treating, preventing,
 PT diagnosing and/or prognosing gastrointestinal disorders, infections,
 PT cancers such as intestinal neoplasms, ulcers.
 XX
 PS Claim 11; SEQ ID NO 84; 179pp; English.
 XX
 CC The invention relates to a colorectal cancer antigen. The antigen is
 CC useful for chromosome identification, chromosome mapping, radiation
 CC hybrid mapping or gene therapy, or as hybridisation probes for
 CC differential identification of the tissues or cell types present in a
 CC biological sample. The antigen is useful for treating, preventing,
 CC diagnosing and/or prognosing gastrointestinal disorders, including
 CC inflammatory diseases and/or conditions, infections, cancers (e.g.
 CC intestinal neoplasms (carcinoid tumour of the small intestine, non-
 CC Hodgkin's lymphoma of the small intestine, small bowel lymphoma)) and
 CC ulcers (e.g. peptic ulcers). The antigen and its nucleic acid are useful
 CC to provide immunological probes for differential identification of the
 CC tissue. The antigen and its nucleic acid are useful for treating,
 CC preventing, diagnosing and/or prognosing diseases, disorders and/or
 CC conditions of the immune system e.g. Bruton's disease, X linked infantile
 CC agammaglobulinemia, severe combined immunodeficiencies, DiGeorge
 CC anomaly, etc. The antigen and its nucleic acid is useful for treating,
 CC preventing and/or diagnosing hyperproliferative disorders (e.g. acute

Query Match 62.7%; Score 37; DB 7; Length 113;
 Best Local Similarity 62.5%; Pred. No. 1.8e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NYVRCRGDD 8
 ||:|:|
 Db 10 NYQCEGED 17

RESULT 37
 ABB11487
 ID ABB11487 standard; peptide; 140 AA.
 XX
 AC ABB11487;
 XX
 DT 11-JAN-2002 (first entry)
 XX
 DE Human secreted protein homologue, SEQ ID NO:1857.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
 KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
 KW cytotatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
 KW antifungal; vulnery; antiulcer.
 XX
 OS Homo sapiens.
 XX
 XX WO200157188-A2.
 PN
 XX

PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US003800.
 XX
 PR 03-FEB-2000; 2000US-00496914.
 PR 27-APR-2000; 2000US-00560875.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 XX Tang YT, Liu C, Drmanac RT;
 PI
 DR WPI; 2001-457740/49.
 DR N-PSDB; ABA08731.
 XX
 PT Human proteins and DNA encoding sequences useful for preventing, treating
 PT or ameliorating a medical condition in a mammalian subject e.g. arthritis
 PT and cancer.
 XX
 PS Claim 20; Page 201; 1963pp; English.
 XX
 CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of identifying the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, and hence
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g. myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g. asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a novel human
 CC polypeptide of the invention
 XX
 SQ Sequence 140 AA;
 Query Match 62.7%; Score 37; DB 4; Length 140;
 Best Local Similarity 85.7%; Pred. No. 2.2e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 4 CRGDDSK 10
 ||:|:|
 Db 80 CRGDDSK 86

RESULT 38
 ABO65272
 ID ABO65272 standard; protein; 275 AA.
 XX

AC ABO65272;
XX
DT 29-JUL-2004 (first entry)
DE Klebsiella pneumoniae polypeptide seqid 11789.
XX
XX Recombinant expression vector; transcription regulatory element;
KW Klebsiella pneumoniae protein; antibacterial; Vaccine.
XX
OS Klebsiella pneumoniae.
XX
PN US6610836-B1.
XX
PD 26-AUG-2003.
XX
PF 27-JAN-2000; 2000US-00489039.
XX
XX 29-JAN-1999; 99US-0117747P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Breton GL, Osborne M;
XX
XX WPI; 2003-895346/82.
DR N-PSDB; ACH98823.
XX
XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
PT preparing a vaccine composition against Klebsiella pneumoniae.
XX
XX Disclosure; SEQ ID NO 11789; 932pp; English.
XX
XX The invention describes a new isolated nucleic acid encoding a Klebsiella
CC pneumoniae polypeptide. Also described are: a recombinant expression
CC vector comprising the nucleic acid, operably linked to a transcription
CC regulatory element; and a cell comprising the recombinant expression
CC vector. The nucleic acid is useful for preparing a vaccine composition
CC against Klebsiella pneumoniae. This is the amino acid sequence of a
CC Klebsiella pneumoniae polypeptide of the invention
XX
XX Sequence 275 AA;
SQ
Query Match 62.7%; Score 37; DB 7; Length 275;
Best Local Similarity 66.7%; Pred. No. 4.3e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 NVRCRGDD 9
Db 32 NWRCGGDS 40
RESULT 39
ABO66647
ID ABO66647 standard; protein; 488 AA.
XX
XX ABO66647;
XX
XX 29-JUL-2004 (first entry)
XX
XX Klebsiella pneumoniae polypeptide seqid 13164.
XX
XX Recombinant expression vector; transcription regulatory element;
KW Klebsiella pneumoniae protein; antibacterial; Vaccine.
XX
XX Klebsiella pneumoniae.
XX
XX US6610836-B1.
PN
PD 26-AUG-2003.
XX
XX 27-JAN-2000; 2000US-00489039.
PF
XX 29-JAN-1999; 99US-0117747P.
XX

PA (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Breton GL, Osborne M;
XX
XX WPI; 2003-895346/82.
DR N-PSDB; ABD00218.
XX
XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
PT preparing a vaccine composition against Klebsiella pneumoniae.
XX
XX Disclosure; SEQ ID NO 13164; 932pp; English.
XX
XX The invention describes a new isolated nucleic acid encoding a Klebsiella
CC pneumoniae polypeptide. Also described are: a recombinant expression
CC vector comprising the nucleic acid, operably linked to a transcription
CC regulatory element; and a cell comprising the recombinant expression
CC vector. The nucleic acid is useful for preparing a vaccine composition
CC against Klebsiella pneumoniae. This is the amino acid sequence of a
CC Klebsiella pneumoniae polypeptide of the invention
XX
XX Sequence 488 AA;
SQ
Query Match 62.7%; Score 37; DB 7; Length 488;
Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 RCRGDD 8
Db 102 RCRGDD 107
RESULT 40
ADD46237
ID ADD46237 standard; protein; 1269 AA.
XX
XX ADD46237;
AC
XX 29-JAN-2004 (first entry)
DT
XX
XX Rat Protein U49055, SEQ ID NO 11912.
DE
XX
XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
XX Rattus norvegicus.
OS
XX WO2003016475-A2.
PN
XX 27-FEB-2003.
PD
XX
XX 14-AUG-2002; 2002WO-US025765.
PF
XX
XX 14-AUG-2001; 2001US-0312147P.
PR
XX 01-NOV-2001; 2001US-0346382P.
PR
XX 26-NOV-2001; 2001US-0333347P.
PR
XX (GEO) GEN HOSPITAL CORP.
PA
XX (FARB) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
PI
XX
XX WPI; 2003-268312/26.
DR
XX GENBANK; U49055.
DR
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC

CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 1269 AA;
 SQ

Query Match 62.7%; Score 37; DB 7; Length 1269;
 Best Local Similarity 85.7%; Pred. No. 1.9e+03;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 YRCRGDD 8
 DB 100 YRCPGDD 106
 ||| |||

RESULT 41
 ADD46239
 ID ADD46239 standard; protein; 1272 AA.
 XX
 AC ADD46239;
 XX
 XX 29-JAN-2004 (first entry)
 XX
 DE Human Protein AB029039, SEQ ID NO 11914.
 XX
 KW Human; pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; chronic constriction injury; CCI;
 KW spared nerve injury; SNI; Chung.
 XX
 XX Homo sapiens.
 OS
 XX WO2003016475-A2.
 PN
 PD 27-FEB-2003.
 XX

PF 14-AUG-2002; 2002WO-US025765.
 XX
 XX 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX
 XX (GEO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 XX
 XX Woolf C, D'urso D, Befort K, Costigan M;
 XX
 XX WPI; 2003-268312/26.
 DR GENBANK; AB029039.
 XX

PT New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 XX
 PS Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a human protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 1272 AA;
 SQ

Query Match 62.7%; Score 37; DB 7; Length 1272;
 Best Local Similarity 85.7%; Pred. No. 2e+03;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 YRCRGDD 8
 DB 100 YRCPGDD 106
 ||| |||

RESULT 42
 AAU49621
 ID AAU49621 standard; protein; 59 AA.
 XX
 AC AAU49621;
 XX
 XX 27-FEB-2002 (first entry)
 DT
 XX
 DE Propionibacterium acnes immunogenic protein #10517.

XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX

OS Propionibacterium acnes.

XX WO200181581-A2.
 PN
 XX 01-NOV-2001.
 PD
 XX 20-APR-2001; 2001WO-US012865.
 PF
 XX 21-APR-2000; 2000US-0199047P.
 PR 02-JUN-2000; 2000US-0208841P.
 PR 07-JUL-2000; 2000US-0216747P.
 XX
 XX (CORI-) CORIXA CORP.
 PA
 XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 XX WPI; 2001-616774/71.

DR N-PSDB; AAS959545.
 XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.
 XX
 XX Example 1; SEQ ID NO 10816; 1069pp; English.
 XX
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 59 AA;
 Query Match 61.0%; Score 36; DB 4; Length 59;
 Best Local Similarity 85.7%; Pred. No. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 YRCRGDD 8
 Db 25 YRCRTDD 31
 |||||
 |||||
 RESULT 43
 ABM46140
 ID ABM46140 standard; protein; 59 AA.
 XX
 AC ABM46140;
 XX
 DT 20-OCT-2003 (first entry)
 XX
 DE Propionibacterium acnes predicted ORF-encoded polypeptide #10816.
 XX
 KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;
 KW immunostimulant; immune response; vaccine.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO2003033515-A1.
 XX
 PD 24-APR-2003.
 XX
 PF 11-OCT-2002; 2002WO-US032727.
 XX
 PR 15-OCT-2001; 2001US-00978825.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Mitcham JL, Skeiky YAM, Persing DH, Bhatia A, Maisonneuve JL;
 PI Zhang Y, Wang S, Jen S, Lodes WJ, Benson DR, Jones R, Carter D;
 PI Barth B, Vallieue-Douglass J;
 XX
 DR WPI; 2003-381789/36.
 DR N-PSDB; ACF64474.
 XX
 PT New Propionibacterium acnes polypeptides and polynucleotides encoding the

PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
 PT or for stimulating an immune response specific for a P. acnes protein.
 XX
 PS Example 1; SEQ ID NO 10816; 1481pp; English.
 XX
 CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
 CC encoding a Propionibacterium acnes protein. The invention also relates to
 CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
 CC immunogenic fragments of P. acnes polypeptides. The invention
 CC additionally encompasses expression vectors and host cells comprising a
 CC polynucleotide of the invention; antibodies against polypeptides of the
 CC polynucleotide of the invention; fusion proteins comprising a polypeptide of the invention; a
 CC method for stimulating an immune response specific for a P. acnes
 CC polypeptide and an isolated T cell population comprising T cells prepared
 CC via this method; a vaccine composition (comprising P. acnes polypeptides,
 CC polynucleotides, antibodies, fusion proteins, T cell populations, or
 CC antigen-presenting cells that express the polypeptide); a method and kit
 CC for detecting or determining the presence or absence of P. acnes in a
 CC patient; and a method for inhibiting the development of P. acnes in a
 CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
 CC proteins, T cell populations or antigen-presenting cells that express the
 CC polypeptides are useful for diagnosing, preventing or treating acne
 CC vulgaris, or for stimulating an immune response specific for a P. acnes
 CC protein. The polynucleotides can also be used as probes or primers for
 CC nucleic acid hybridisation. The vaccine composition is useful for the
 CC stimulation of an immune response against P. acnes, or for treating acne,
 CC and the kit is useful for performing a diagnostic assay. The present
 CC sequence represents a polypeptide predicted to be encoded by an ORF (open
 CC reading frame) contained within the P. acnes polynucleotides of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 59 AA;
 Query Match 61.0%; Score 36; DB 6; Length 59;
 Best Local Similarity 85.7%; Pred. No. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 YRCRGDD 8
 Db 25 YRCRTDD 31
 |||||
 |||||
 RESULT 44
 ABM69291
 ID ABM69291 standard; protein; 405 AA.
 XX
 AC ABM69291;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Photorhabdus luminescens protein sequence #2388.
 XX
 KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
 KW detection; food; gene expression; plant; animal; microorganism; toxin;
 KW antibiotic; biopesticide; virulence factor; disease model; plague;
 KW whooping cough.
 XX
 OS Photorhabdus luminescens.
 XX
 PN WO200294867-A2.
 XX
 PD 28-NOV-2002.
 XX
 PF 07-FEB-2002; 2002WO-IB003040.
 XX
 PR 07-FEB-2001; 2001FR-00001659.
 XX
 PA (INSP) INST PASTEUR.
 PA (CNRS) CNRS CENT NAT RECH SCI.
 XX
 PI Duchaud E, Taurit S, Glaser P, Frangeul L, Kunst F, Danchin A;

PI Buchrieser C;
 XX WPI; 2003-148459/14.
 XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
 XX useful e.g. as therapeutic antimicrobials and agricultural pesticides.
 XX
 XX Claim 2; SEQ ID NO 2388; 1205pp; French.
 XX
 XX The invention relates to the isolation of genes and their encoded
 CC proteins from Photorhabdus luminescens. The isolated sequences are
 CC sources of probes and primers for detecting the genome of P. luminescens
 CC and related species; to study polymorphisms; for gene analysis and for
 CC detection/amplification of the genes. Antibodies (Ab) raised against the
 CC polypeptides encoded by the genes are used for detection/identification
 CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
 CC carry a gene-containing vector are used to select compounds that
 CC modulate, regulate, induce or inhibit expression of the genes in plants,
 CC animals or microorganisms other than P. luminescens and are able to alter
 CC response or sensitivity to toxins and antibiotics produced by P.
 CC luminescens. Cells transformed to express the genes are useful for
 CC recombinant production of the proteins, particularly toxins and
 CC antibacterials useful as insecticides, bactericides and fungicides. The
 CC genes, proteins, vectors containing the genes and Ab are also useful
 CC therapeutically (to treat microbial infection by bacteria or fungi that
 CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
 CC biopesticides. Other uses of the genes and the proteins are as virulence
 CC factors and for identifying targets of human diseases for which P.
 CC luminescens is a model (particularly plague and whooping cough). This
 CC sequence represents one of the isolated P. luminescens proteins
 XX
 XX Sequence 405 AA;
 XX
 XX Query Match 61.0%; Score 36; DB 6; Length 405;
 XX Best Local Similarity 71.4%; Pred. No. 9.2e+02;
 XX Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 2 YRCRGDD 8
 DB :|||||:
 DB 39 FRCRGDE 45
 RESULT 45
 ADN20343
 ID ADN20343 standard; protein; 466 AA.
 XX
 XX AC ADN20343;
 XX
 XX DT 02-DEC-2004 (first entry)
 XX
 XX DE Bacterial polypeptide #2996.
 XX
 XX KW Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.
 XX
 XX OS Bacteria.
 XX
 XX XX US2003233675-A1.
 XX
 XX PD 18-DEC-2003.
 XX
 XX PF 20-FEB-2003; 2003US-00369493.
 XX
 XX PR 21-FEB-2002; 2002US-0360039P.
 XX
 XX PA (CAOY/) CAO Y.
 XX (HINK/) HINKLE G J.
 XX (SLAT/) SLATER S C.

PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX
 XX DR WPI; 2004-061375/06.
 XX
 XX PT New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX
 XX PS Claim 1; SEQ ID NO 2996; 122pp; English.
 XX
 XX CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 XX SQ Sequence 466 AA;
 XX
 XX Query Match 61.0%; Score 36; DB 8; Length 466;
 XX Best Local Similarity 71.4%; Pred. No. 1.1e+03;
 XX Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 2 YRCRGDD 8
 DB :|||||:
 DB 214 FRCRGED 220
 RESULT 46
 ADN20289
 ID ADN20289 standard; protein; 470 AA.
 XX
 XX AC ADN20289;
 XX
 XX DT 02-DEC-2004 (first entry)
 XX
 XX DE Bacterial polypeptide #2942.
 XX
 XX KW Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.
 XX
 XX OS Bacteria.
 XX
 XX XX US2003233675-A1.
 XX
 XX PD 18-DEC-2003.
 XX
 XX XX

PF 20-FEB-2003; 2003US-00369493.

PR 21-FEB-2002; 2002US-0360039P.

XX (CAOY/) CAO Y.

PA (HINK/) HINKLE G J.

PA (SLAT/) SLATER S C.

PA (CHEN/) CHEN X.

PA (GOLD/) GOLDMAN B S.

XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI; 2004-061375/06.

XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.

XX Claim 1; SEQ ID NO 2942; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

SQ Sequence 470 AA;

Query Match 61.0%; Score 36; DB 8; Length 470;

Best Local Similarity 71.4%; Pred. No. 1.1e+03;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YRCRGDD 8

Db 214 FRCRGED 220

RESULT 47

ADM04173

ID ADM04173 standard; protein; 485 AA.

XX ADM04173;

XX 20-MAY-2004 (first entry)

DE Human protein of the invention SEQ ID NO:2858.

XX human; gene therapy; diagnostic marker; pharmaceutical.

OS Homo sapiens.

XX EP1347046-A1.

XX 24-SEP-2003.

PD

XX 12-APR-2002; 2002EP-00008400.

XX 22-MAR-2002; 2002JP-00137785.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;

PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;

PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;

XX WPI; 2003-723558/69.

DR N-PSDB; ADM01730.

XX New polynucleotides and polypeptides are useful in gene therapy, for
PT developing a diagnostic marker or medicines for regulating their
PT expression and activity, or as a target of gene therapy.

XX Claim 1; SEQ ID NO 2858; 305pp; English.

XX The invention relates to a novel human polynucleotide and the encoded
CC polypeptide. A polynucleotide of the invention may have a use in gene
CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
CC as a primer for synthesizing the polynucleotide or as a probe for
CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are
CC useful in gene therapy, for developing a diagnostic marker or medicines
CC for regulating their expression and activity, or as a target of gene
CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
CC are useful as pharmaceutical agents. The present sequence represents a
CC protein sequence of the invention.

SQ Sequence 485 AA;

Query Match 61.0%; Score 36; DB 7; Length 485;

Best Local Similarity 60.0%; Pred. No. 1.1e+03;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NVRCRGDDSK 10

Db 469 NWRCPGTUSE 478

RESULT 48

AAAY79209

ID AAY79209 standard; protein; 498 AA.

XX AAY79209;

XX 19-JUN-2000 (first entry)

XX Human transferase TRNSFS-1.

XX Transferase; TRNSFS-1; human; antitumour; cancer;

XX gastrointestinal disorder; developmental disorder; genetic disorder;

XX neurological disorder; reproductive disorder; smooth muscle disorder;

XX immunological disorder; inflammation; diagnosis; therapy;

XX myristoyl CoA:protein N-myristoyltransferase.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 64

FT /note= "potential O-phosphorylation"

FT Modified-site 68

FT /note= "potential O-phosphorylation"

FT Peptide 150..228

FT /note= "myristoyl-CoA protein N-myristoyltransferase signature"

FT Modified-site 166

FT /note= "potential O-phosphorylation"

FT Modified-site 220

FT /note= "potential O-phosphorylation"

FT Peptide 240..285

CC lymphatic growth agent may also be used in manufacturing a medicament for
 CC the treatment of hereditary lymphoedema resulting from a mutation in a
 CC LEC gene or of other diseases involving the lymphatic vessels, such as
 CC various inflammatory diseases and cancer metastasis via the lymphatic
 CC system. The present sequence is that of a human LEC/BEC differentially
 CC expressed protein which is related to the method of the invention. Note:
 CC This sequence does not appear in the specification but was obtained by
 CC the indexer using the source data given in table 14 of the specification.

XX Sequence 498 AA;

Query Match 61.0%; Score 36; DB 7; Length 498;

Best Local Similarity 60.0%; Pred. NO. 1.1e+03;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10

Db |:||| |::

482 NWRCPGTDSE 491

RESULT 50

ADP23807

ID ADP23807 standard; protein; 498 AA.

XX AC

ADP23807;

DT 18-NOV-2004 (first entry)

XX PRO polypeptide SEQ ID NO:985.

DE PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;
 KW osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;
 KW antiasthmatic; hepatotropic; respiratory; gene therapy; immune system.

XX Unidentified.

XX WO2004041170-A2.

XX 21-MAY-2004.

XX 30-OCT-2003; 2003WO-US034312.

XX 01-NOV-2002; 2002US-0423394P.

XX (GETH) GENENTECH INC.

XX Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;

PI Wu TD;

XX WPI; 2004-419628/39.

DR N-PSDB; ADP23806.

XX New PRO polypeptides and polynucleotides, useful for treating e.g.
 PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated
 PT renal disease, or demyelinating diseases of the central or peripheral
 PT nervous system.

PS Claim 7; SEQ ID NO 985; 2940pp; English.

XX The invention relates to a novel isolated nucleic acid and the PRO
 CC polypeptide encoded by it. A protein of the invention has
 CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
 CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
 CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide
 CC of the invention may have a use in gene therapy. The PRO polypeptide, its
 CC agonist, antagonist, or antibody that specifically binds to the
 CC polypeptide is useful for treating an immune related disorder such as
 CC systemic lupus erythematous, rheumatoid arthritis, osteoarthritis,
 CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
 CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
 CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
 CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
 CC disease, a demyelinating disease of the central or peripheral nervous

CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
 CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
 CC disease, infectious or autoimmune chronic active hepatitis, primary
 CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
 CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
 CC disease, an autoimmune or immune-mediated skin disease, a bullous skin
 CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
 CC disease, asthma, allergic rhinitis, atopic dermatitis, food
 CC hypersensitivity, urticaria, an immunologic disease of the lung,
 CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
 CC pneumonitis, a transplantation associated disease, graft rejection or
 CC graft-versus-host disease. The present sequence represents a PRO protein
 CC of the invention.

XX Sequence 498 AA;

Query Match 61.0%; Score 36; DB 8; Length 498;

Best Local Similarity 60.0%; Pred. NO. 1.1e+03;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NYRCRGDDSK 10

Db |:||| |::

482 NWRCPGTDSE 491

Search completed: September 7, 2005, 19:54:56

Job time : 86.093 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2005, 19:48:00 ; Search time 132.558 Seconds
(without alignments)
59.419 Million cell updates/sec

Title: US-10-812-238B-2

Perfect score: 109

Sequence: 1 EGVQNYRCRGDDSKVQEAR 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1774312 seqs, 393823214 residues

Total number of hits satisfying chosen parameters: 1774312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	109	100.0	20	17 US-10-812-238A-2	Sequence 2, Appli
2	109	100.0	311	16 US-10-655-601-2	Sequence 2, Appli
3	109	100.0	311	17 US-10-812-238A-13	Sequence 13, Appli
4	73	67.0	13	17 US-10-812-238A-23	Sequence 23, Appli
5	59	54.1	10	17 US-10-812-238A-20	Sequence 20, Appli
6	59	54.1	221	15 US-10-287-226-346	Sequence 346, App
7	59	54.1	221	15 US-10-287-226-348	Sequence 348, App
8	59	54.1	284	18 US-10-764-425-157	Sequence 157, App
9	59	54.1	285	16 US-10-655-601-1	Sequence 1, Appli
10	59	54.1	289	16 US-10-643-795A-136	Sequence 136, App
11	59	54.1	289	17 US-10-948-518-136	Sequence 136, App

55	50.5	10	17	US-10-812-238A-22	Sequence 22, Appli
53	48.6	10	17	US-10-812-238A-21	Sequence 21, Appli
49	45.0	847	15	US-10-369-493-13157	Sequence 13157, A
15	45.0	1327	15	US-10-369-493-13172	Sequence 13172, A
46	42.2	74	16	US-10-425-115-262985	Sequence 262985,
17	42.2	304	15	US-10-424-599-228315	Sequence 228315,
18	41.3	135	15	US-10-316-194-17	Sequence 17, Appli
19	41.3	135	17	US-10-732-180-17	Sequence 17, Appli
20	41.3	822	9	US-09-147-947-6	Sequence 6, Appli
21	41.3	875	16	US-10-757-262-34	Sequence 34, Appli
22	41.3	875	17	US-10-723-860-4046	Sequence 4046, Ap
23	41.3	875	17	US-10-843-299-2	Sequence 2, Appli
24	40.4	99	16	US-10-437-963-108989	Sequence 108989,
25	40.4	191	15	US-10-114-270-104	Sequence 104, App
26	40.4	485	15	US-10-108-260A-2858	Sequence 2858, Ap
27	40.4	498	14	US-10-427-631-1	Sequence 1, Appli
28	40.4	1100	15	US-10-369-493-13176	Sequence 13176, A
29	40.4	1907	9	US-09-832-292-39	Sequence 39, Appli
30	40.4	3993	16	US-10-437-963-168323	Sequence 168323,
31	43.5	130	15	US-10-424-599-202541	Sequence 202541,
32	43	193	16	US-10-767-701-36687	Sequence 36687, A
33	43	237	15	US-10-291-172-610	Sequence 610, App
34	43	237	15	US-10-221-278-610	Sequence 610, App
35	43	363	14	US-10-156-761-13203	Sequence 13203, A
36	43	421	15	US-10-210-281-12	Sequence 12, Appli
37	43	424	15	US-10-210-281-10	Sequence 10, Appli
38	43	428	14	US-10-157-031-139	Sequence 139, App
39	43	489	15	US-10-108-260A-2988	Sequence 2988, Ap
40	43	505	15	US-10-369-493-21996	Sequence 21996, A
41	43	557	9	US-09-815-242-12804	Sequence 12804, A
42	43	565	15	US-10-282-122A-43871	Sequence 43871, A
43	43	565	17	US-10-857-625-554	Sequence 554, App
44	43	566	14	US-10-282-287-4	Sequence 4, Appli
45	42	130	15	US-10-282-122A-78271	Sequence 78271, A
46	42	137	15	US-10-104-047-3193	Sequence 3193, Ap
47	42	173	16	US-10-767-701-33785	Sequence 33785, A
48	42	208	16	US-10-437-963-135805	Sequence 135805,
49	42	305	18	US-10-898-775-28	Sequence 28, Appli
50	42	38.5	16	US-10-437-963-135808	Sequence 135808,
51	42	38.5	15	US-10-282-122A-67892	Sequence 67892, A
52	42	1166	14	US-10-032-585-7638	Sequence 7638, Ap
53	41	37.6	56	US-10-424-599-268486	Sequence 268486,
54	41	37.6	119	US-10-425-115-337109	Sequence 337109,
55	41	37.6	130	US-10-282-122A-68446	Sequence 68446, A
56	41	37.6	201	US-10-434-817-12	Sequence 12, Appli
57	41	37.6	201	US-11-086-903-12	Sequence 12, Appli
58	41	37.6	236	US-10-434-817-7	Sequence 7, Appli
59	41	37.6	236	US-11-086-903-7	Sequence 7, Appli
60	41	37.6	255	US-10-434-817-2	Sequence 2, Appli
61	41	37.6	255	US-11-086-903-2	Sequence 2, Appli
62	41	37.6	306	US-10-767-701-39559	Sequence 39559, A
63	41	37.6	396	US-10-065-133A-92	Sequence 92, Appli
64	41	37.6	396	US-10-065-133A-95	Sequence 95, Appli
65	41	37.6	396	US-10-734-373-92	Sequence 92, Appli
66	41	37.6	396	US-10-734-373-95	Sequence 95, Appli
67	41	37.6	396	US-10-181-585B-64	Sequence 64, Appli
68	41	37.6	396	US-10-181-585B-67	Sequence 67, Appli
69	41	37.6	752	US-10-381-530-28	Sequence 28, Appli
70	41	37.6	757	US-10-065-133A-107	Sequence 107, App
71	41	37.6	757	US-10-734-373-107	Sequence 107, App
72	41	37.6	757	US-10-181-585B-72	Sequence 72, Appli
73	41	37.6	1150	US-10-437-963-147798	Sequence 147798,
74	40.5	37.2	242	US-10-425-115-307140	Sequence 307140,
75	40.5	37.2	255	US-10-437-963-116185	Sequence 116185,
76	40.5	37.2	353	US-10-437-963-178169	Sequence 178169,
77	40.5	37.2	373	US-09-738-630-80	Sequence 80, Appli
78	40	36.7	42	US-10-425-115-328990	Sequence 328990,
79	40	36.7	49	US-10-424-599-152636	Sequence 152636,
80	40	36.7	51	US-10-437-963-204484	Sequence 204484,
81	40	36.7	79	US-10-437-963-131923	Sequence 131923,
82	40	36.7	84	US-10-437-963-154748	Sequence 154748,
83	40	36.7	94	US-10-289-762-1290	Sequence 1290, Ap
84	40	36.7	102	US-10-424-599-260014	Sequence 260014,

85	40	36.7	111	15	US-10-424-599-254922	Sequence 254922,	158	39.5	36.2	211	15	US-10-341-434-77	Sequence 77, Appl
86	40	36.7	112	15	US-10-424-599-275302	Sequence 275302,	159	39.5	36.2	211	15	US-10-236-031B-14	Sequence 14, Appl
87	40	36.7	113	15	US-10-282-122A-50965	Sequence 50965, A	160	39.5	36.2	211	15	US-10-072-012-335	Sequence 335, App
88	40	36.7	112	15	US-10-282-122A-76429	Sequence 76429, A	161	39.5	36.2	211	15	US-10-072-012-336	Sequence 336, App
89	40	36.7	140	15	US-10-276-774-1857	Sequence 1857, Ap	162	39.5	36.2	211	15	US-10-344-487-13	Sequence 13, Appl
90	40	36.7	147	15	US-10-424-599-269282	Sequence 269282,	163	39.5	36.2	211	15	US-10-601-953-896	Sequence 896, App
91	40	36.7	231	15	US-10-424-599-239306	Sequence 239306,	164	39.5	36.2	211	15	US-10-106-698-6273	Sequence 6273, Ap
92	40	36.7	226	18	US-10-724-972A-5260	Sequence 5260, Ap	165	39.5	36.2	247	14	US-10-264-049-3053	Sequence 3053, Ap
93	40	36.7	259	15	US-10-424-599-222209	Sequence 222209,	166	39.5	36.2	286	15	US-10-282-122A-69355	Sequence 69355, A
94	40	36.7	286	15	US-10-424-599-173051	Sequence 173051,	167	39.5	36.2	453	15	US-10-369-493-1571	Sequence 4172, Ap
95	40	36.7	301	15	US-10-369-493-17538	Sequence 17538, A	168	39	35.8	43	16	US-10-437-963-17113	Sequence 157113,
96	40	36.7	301	15	US-10-369-493-16124	Sequence 16124, A	169	39	35.8	51	16	US-10-435-115-309902	Sequence 309902,
97	40	36.7	302	15	US-10-369-493-15371	Sequence 15371, A	170	39	35.8	74	9	US-09-764-877-2043	Sequence 2043, Ap
98	40	36.7	371	15	US-10-425-114-42592	Sequence 42592, A	171	39	35.8	74	15	US-10-242-515-2043	Sequence 2043, Ap
99	40	36.7	396	14	US-10-065-133A-86	Sequence 86, Appl	172	39	35.8	77	16	US-10-437-963-170560	Sequence 170560,
100	40	36.7	386	14	US-10-065-133A-89	Sequence 89, Appl	173	39	35.8	99	16	US-10-437-963-149116	Sequence 149116,
101	40	36.7	396	16	US-10-734-373-86	Sequence 86, Appl	174	39	35.8	115	15	US-10-424-599-165388	Sequence 165388,
102	40	36.7	396	16	US-10-734-373-89	Sequence 89, Appl	175	39	35.8	117	15	US-10-424-599-249919	Sequence 249919,
103	40	36.7	396	18	US-10-181-585B-54	Sequence 54, Appl	176	39	35.8	122	11	US-09-864-408A-8608	Sequence 8608, Ap
104	40	36.7	396	18	US-10-181-585B-52	Sequence 62, Appl	177	39	35.8	127	16	US-10-467-253-9	Sequence 9, Appli
105	40	36.7	400	15	US-10-282-122A-43293	Sequence 43293, A	178	39	35.8	131	15	US-10-282-122A-50058	Sequence 50058, A
106	40	36.7	400	15	US-10-282-122A-59506	Sequence 59506, A	179	39	35.8	132	9	US-09-738-626-4098	Sequence 4098, Ap
107	40	36.7	400	15	US-10-282-122A-73186	Sequence 73186, A	180	39	35.8	132	16	US-10-494-674-16	Sequence 16, Appl
108	40	36.7	400	15	US-10-282-122A-75975	Sequence 75975, A	181	39	35.8	161	16	US-10-767-701-54578	Sequence 54578, A
109	40	36.7	401	14	US-10-156-761-9009	Sequence 9009, Ap	182	39	35.8	173	9	US-09-205-658-216	Sequence 216, App
110	40	36.7	421	16	US-10-437-963-204884	Sequence 204884,	183	39	35.8	173	10	US-09-963-693-216	Sequence 216, App
111	40	36.7	422	16	US-10-739-930-6688	Sequence 6688, Ap	184	39	35.8	185	15	US-10-282-122A-63885	Sequence 63885, A
112	40	36.7	434	14	US-10-167-127-13	Sequence 13, Appl	185	39	35.8	214	14	US-10-002-631C-250	Sequence 250, App
113	40	36.7	440	15	US-10-369-493-3576	Sequence 3576, Ap	186	39	35.8	234	15	US-10-424-599-179999	Sequence 179999,
114	40	36.7	482	16	US-10-425-115-210789	Sequence 210789,	187	39	35.8	253	15	US-10-282-122A-63937	Sequence 63937, A
115	40	36.7	493	16	US-10-437-963-131895	Sequence 131895,	188	39	35.8	265	16	US-10-437-963-178538	Sequence 178538,
116	40	36.7	497	16	US-10-437-963-111779	Sequence 111779,	189	39	35.8	297	15	US-10-282-122A-51933	Sequence 51933, A
117	40	36.7	531	15	US-10-259-194A-150	Sequence 150, App	190	39	35.8	301	9	US-09-815-242-11200	Sequence 11200, A
118	40	36.7	531	16	US-10-425-115-358488	Sequence 358488,	191	39	35.8	325	15	US-10-104-047-3498	Sequence 3498, Ap
119	40	36.7	537	16	US-10-437-963-179921	Sequence 179921,	192	39	35.8	374	16	US-10-425-115-225918	Sequence 225918,
120	40	36.7	543	15	US-10-425-114-67816	Sequence 67816, A	193	39	35.8	402	15	US-10-104-047-3637	Sequence 3637, Ap
121	40	36.7	543	16	US-10-767-701-46229	Sequence 46229, A	194	39	35.8	435	15	US-10-425-114-43540	Sequence 43540, A
122	40	36.7	662	15	US-10-424-599-208782	Sequence 208782,	195	39	35.8	447	13	US-10-047-676A-10	Sequence 10, Appl
123	40	36.7	669	16	US-10-425-115-215599	Sequence 215599,	196	39	35.8	447	16	US-10-790-914-10	Sequence 10, Appl
124	40	36.7	730	14	US-10-067-632-58	Sequence 58, Appl	197	39	35.8	461	15	US-10-417-700A-21	Sequence 21, Appl
125	40	36.7	757	14	US-10-073-377-23	Sequence 23, Appl	198	39	35.8	462	17	US-10-732-923-6182	Sequence 6182, Ap
126	40	36.7	757	14	US-10-073-377-25	Sequence 25, Appl	199	39	35.8	462	17	US-10-732-923-6187	Sequence 6187, Ap
127	40	36.7	757	14	US-10-073-377-27	Sequence 27, Appl	200	39	35.8	466	15	US-10-369-493-2996	Sequence 2996, Ap
128	40	36.7	757	14	US-10-073-377-29	Sequence 29, Appl	201	39	35.8	470	15	US-10-369-493-2942	Sequence 2942, Ap
129	40	36.7	757	14	US-10-073-377-31	Sequence 31, Appl	202	39	35.8	487	15	US-10-424-599-269280	Sequence 269280,
130	40	36.7	757	14	US-10-073-377-33	Sequence 33, Appl	203	39	35.8	487	15	US-10-425-114-51705	Sequence 51705, A
131	40	36.7	757	14	US-10-073-377-35	Sequence 35, Appl	204	39	35.8	540	16	US-10-437-963-111986	Sequence 111986,
132	40	36.7	757	14	US-10-073-377-37	Sequence 37, Appl	205	39	35.8	550	15	US-10-108-260A-4334	Sequence 4334, Ap
133	40	36.7	757	14	US-10-073-377-39	Sequence 39, Appl	206	39	35.8	557	9	US-09-815-242-12165	Sequence 12165, A
134	40	36.7	757	14	US-10-073-377-41	Sequence 41, Appl	207	39	35.8	563	9	US-09-815-242-5464	Sequence 5464, Ap
135	40	36.7	757	14	US-10-073-377-43	Sequence 43, Appl	208	39	35.8	565	16	US-10-408-765A-1726	Sequence 1726, Ap
136	40	36.7	757	14	US-10-073-377-45	Sequence 45, Appl	209	39	35.8	572	15	US-10-425-114-43039	Sequence 43039, A
137	40	36.7	757	14	US-10-073-377-47	Sequence 47, Appl	210	39	35.8	602	15	US-10-094-749-3150	Sequence 3150, Ap
138	40	36.7	757	14	US-10-065-133A-104	Sequence 104, App	211	39	35.8	615	17	US-10-741-600-1522	Sequence 1522, Ap
139	40	36.7	757	16	US-10-381-530-10	Sequence 10, Appl	212	39	35.8	631	15	US-10-424-599-266585	Sequence 266585,
140	40	36.7	757	16	US-10-734-373-104	Sequence 104, App	213	39	35.8	654	15	US-10-369-493-5125	Sequence 5125, Ap
141	40	36.7	757	18	US-10-181-585B-69	Sequence 69, Appl	214	39	35.8	662	17	US-10-741-600-1523	Sequence 1523, Ap
142	40	36.7	823	14	US-10-167-127-15	Sequence 15, Appl	215	39	35.8	662	17	US-10-741-600-1524	Sequence 1524, Ap
143	40	36.7	891	17	US-10-732-923-14985	Sequence 14985, A	216	39	35.8	670	14	US-10-156-761-14107	Sequence 14107, A
144	40	36.7	967	17	US-10-732-923-14984	Sequence 14984, A	217	39	35.8	690	15	US-10-369-493-22730	Sequence 22730,
145	40	36.7	1896	17	US-10-732-923-1758	Sequence 1758, A	218	39	35.8	693	16	US-10-437-963-149803	Sequence 149803,
146	40	36.7	5560	15	US-10-263-929-146	Sequence 142, App	219	39	35.8	694	16	US-10-425-115-210914	Sequence 210914,
147	40	36.7	8523	18	US-10-840-512-162	Sequence 162, App	220	39	35.8	702	17	US-10-741-600-1526	Sequence 1526, Ap
148	39.5	36.2	77	16	US-10-425-115-230658	Sequence 290658,	221	39	35.8	754	17	US-10-741-600-1525	Sequence 1525, Ap
149	39.5	36.2	77	16	US-10-425-115-230658	Sequence 290658,	222	39	35.8	801	9	US-09-815-242-5518	Sequence 5518, Ap
150	39.5	36.2	78	16	US-10-437-963-139394	Sequence 139394,	223	39	35.8	801	9	US-09-815-242-12519	Sequence 12519, A
151	39.5	36.2	126	14	US-10-002-631C-46	Sequence 46, Appl	224	39	35.8	801	15	US-10-282-122A-44225	Sequence 44225, A
152	39.5	36.2	137	10	US-09-866-050A-174	Sequence 174, Appl	225	39	35.8	809	15	US-10-282-122A-61637	Sequence 61637, A
153	39.5	36.2	194	9	US-09-925-301-999	Sequence 999, App	226	39	35.8	898	15	US-10-257-174-43	Sequence 43, Appl
154	39.5	36.2	210	9	US-09-935-390A-36	Sequence 36, Appl	227	39	35.8	912	13	US-10-047-542-75	Sequence 75, Appl
155	39.5	36.2	211	10	US-09-471-749-7	Sequence 7, Appli	228	39	35.8	917	13	US-10-047-542-76	Sequence 76, Appl
156	39.5	36.2	211	14	US-10-176-847-8	Sequence 84, Appl	229	39	35.8	917	13	US-10-047-542-87	Sequence 87, Appl
157	39.5	36.2	211	14	US-10-068-486A-2	Sequence 2, Appli	230	39	35.8	917	14	US-10-025-524-2	Sequence 2, Appli

231	39	35.8	924	13	US-10-047-542-67	Sequence 67, Appl	304	34.9	347	10	US-09-907-613-148	Sequence 148, App
232	39	35.8	924	14	US-10-025-524-28	Sequence 28, Appl	305	34.9	347	10	US-09-907-942-148	Sequence 148, App
233	39	35.8	924	16	US-10-723-860-1976	Sequence 1976, Ap	306	34.9	347	10	US-09-904-859-148	Sequence 148, App
234	39	35.8	1004	15	US-10-114-270-32	Sequence 32, Appl	307	34.9	347	10	US-09-909-204-148	Sequence 148, App
235	39	35.8	1229	16	US-10-437-963-133251	Sequence 133251,	308	34.9	347	10	US-09-904-820-148	Sequence 148, App
236	39	35.8	2742	16	US-10-437-963-139002	Sequence 139002,	309	34.9	347	10	US-09-904-786-148	Sequence 148, App
237	39	35.8	3389	15	US-10-016-248-47	Sequence 47, Appl	310	34.9	347	10	US-09-906-646-148	Sequence 148, App
238	39	35.8	3389	16	US-10-408-765A-2286	Sequence 2286, Ap	311	34.9	347	10	US-09-906-700-148	Sequence 148, App
239	39	35.8	3508	15	US-10-016-248-46	Sequence 46, Appl	312	34.9	347	10	US-09-903-786-148	Sequence 148, App
240	38.5	35.3	89	14	US-10-029-386-29445	Sequence 29445, A	313	34.9	347	10	US-09-903-903-148	Sequence 148, App
241	38.5	35.3	114	9	US-09-738-626-6692	Sequence 6692, Ap	314	34.9	347	10	US-09-903-749A-148	Sequence 148, App
242	38.5	35.3	159	15	US-10-108-260A-4046	Sequence 4046, Ap	315	34.9	347	10	US-09-904-119-148	Sequence 148, App
243	38.5	35.3	256	17	US-10-732-923-8967	Sequence 8967, Ap	316	34.9	347	10	US-09-904-956-148	Sequence 148, App
244	38.5	35.3	274	18	US-10-491-467-49	Sequence 49, Appl	317	34.9	347	10	US-09-903-736-148	Sequence 148, App
245	38.5	35.3	553	16	US-10-425-115-352266	Sequence 352266,	318	34.9	347	10	US-09-907-794-148	Sequence 148, App
246	38.5	35.3	645	14	US-10-156-763-12752	Sequence 12752, A	319	34.9	347	10	US-09-903-943-148	Sequence 148, App
247	38.5	35.3	870	16	US-10-437-963-116415	Sequence 116415,	320	34.9	347	10	US-09-904-462-148	Sequence 148, App
248	38.5	35.3	897	15	US-10-259-194A-304	Sequence 304, App	321	34.9	347	10	US-09-907-925-148	Sequence 148, App
249	38.5	35.3	974	16	US-10-437-963-150927	Sequence 150927,	322	34.9	347	10	US-09-902-632-148	Sequence 148, App
250	38.5	35.3	993	16	US-10-437-963-182980	Sequence 182980,	323	34.9	347	10	US-09-903-520-148	Sequence 148, App
251	38.5	35.3	1081	16	US-10-437-963-178760	Sequence 178760,	324	34.9	347	10	US-09-905-056-148	Sequence 148, App
252	38.5	35.3	1275	15	US-10-282-122A-53272	Sequence 53272, A	325	34.9	347	10	US-09-905-064-148	Sequence 148, App
253	38.5	35.3	1287	18	US-10-756-149-5772	Sequence 5772, Ap	326	34.9	347	10	US-09-904-553-148	Sequence 148, App
254	38.5	35.3	1561	17	US-10-805-684-45	Sequence 45, Appl	327	34.9	347	10	US-09-905-381-148	Sequence 148, App
255	38.5	35.3	1893	17	US-10-479-764-15	Sequence 15, Appl	328	34.9	347	10	US-09-904-485-148	Sequence 148, App
256	38	34.9	22	9	US-09-864-761-34955	Sequence 34955, A	329	34.9	347	10	US-09-905-348-148	Sequence 148, App
257	38	34.9	43	15	US-10-424-599-238960	Sequence 238960,	330	34.9	347	10	US-09-905-088-148	Sequence 148, App
258	38	34.9	71	15	US-10-424-599-283295	Sequence 283295,	331	34.9	347	10	US-09-907-575-148	Sequence 148, App
259	38	34.9	77	15	US-10-264-049-3870	Sequence 3870, Ap	332	34.9	347	10	US-09-905-075-148	Sequence 148, App
260	38	34.9	85	16	US-10-437-963-124176	Sequence 124176,	333	34.9	347	10	US-09-902-759-148	Sequence 148, App
261	38	34.9	94	16	US-10-425-115-303573	Sequence 303573,	334	34.9	347	10	US-09-903-634-148	Sequence 148, App
262	38	34.9	96	16	US-10-437-963-201926	Sequence 201926,	335	34.9	347	10	US-09-902-713-148	Sequence 148, App
263	38	34.9	108	15	US-10-108-260A-3952	Sequence 3952, Ap	336	34.9	347	10	US-09-907-979-148	Sequence 148, App
264	38	34.9	116	15	US-10-282-122A-51355	Sequence 51355, A	337	34.9	347	10	US-09-902-615-148	Sequence 148, App
265	38	34.9	129	16	US-10-451-467A-214	Sequence 214, App	338	34.9	347	10	US-09-284-320-4	Sequence 4, Appli
266	38	34.9	136	14	US-10-062-831-131	Sequence 131, App	339	34.9	347	10	US-09-903-925-148	Sequence 148, App
267	38	34.9	136	14	US-10-062-598-131	Sequence 131, App	340	34.9	347	10	US-09-906-760A-148	Sequence 148, App
268	38	34.9	151	15	US-10-074-978A-404	Sequence 404, App	341	34.9	347	10	US-09-903-823-148	Sequence 148, App
269	38	34.9	151	16	US-10-425-115-269918	Sequence 269918,	342	34.9	347	10	US-09-907-652-148	Sequence 148, App
270	38	34.9	152	17	US-10-732-923-22004	Sequence 22004, A	343	34.9	347	10	US-09-902-572A-148	Sequence 148, App
271	38	34.9	159	16	US-10-425-115-346856	Sequence 346856,	344	34.9	347	10	US-09-902-979-148	Sequence 148, App
272	38	34.9	164	9	US-09-863-824-4	Sequence 4, Appli	345	34.9	347	10	US-09-905-125-148	Sequence 148, App
273	38	34.9	164	18	US-10-950-177-4	Sequence 4, Appli	346	34.9	347	10	US-09-906-815A-148	Sequence 148, App
274	38	34.9	168	16	US-10-767-701-54860	Sequence 54860, A	347	34.9	347	10	US-09-905-449-148	Sequence 148, App
275	38	34.9	172	15	US-10-282-122A-54254	Sequence 54254, A	348	34.9	347	10	US-09-903-806-148	Sequence 148, App
276	38	34.9	182	17	US-10-491-823-96	Sequence 96, Appl	349	34.9	347	10	US-09-904-992-148	Sequence 148, App
277	38	34.9	206	13	US-10-087-193-1548	Sequence 14778, Ap	350	34.9	347	10	US-09-904-838-148	Sequence 148, App
278	38	34.9	212	16	US-10-437-963-147771	Sequence 147771,	351	34.9	347	10	US-09-906-777-148	Sequence 148, App
279	38	34.9	228	16	US-10-437-963-143729	Sequence 143729,	352	34.9	347	10	US-09-903-603A-148	Sequence 148, App
280	38	34.9	240	16	US-10-490-318-15	Sequence 15, Appl	353	34.9	347	10	US-09-904-532-148	Sequence 148, App
281	38	34.9	250	15	US-10-282-122A-58270	Sequence 58270, A	354	34.9	347	10	US-09-904-766-148	Sequence 148, App
282	38	34.9	268	15	US-10-424-599-152408	Sequence 152408,	355	34.9	347	10	US-09-904-920A-148	Sequence 148, App
283	38	34.9	303	14	US-10-032-201B-268	Sequence 268, App	356	34.9	347	10	US-09-904-877A-148	Sequence 148, App
284	38	34.9	303	17	US-10-472-928-2968	Sequence 2968, Ap	357	34.9	347	10	US-09-903-562-148	Sequence 148, App
285	38	34.9	305	16	US-10-425-115-283662	Sequence 283662,	358	34.9	347	10	US-09-906-618-148	Sequence 148, App
286	38	34.9	307	16	US-10-322-281-690	Sequence 690, App	359	34.9	347	10	US-09-907-728-148	Sequence 148, App
287	38	34.9	318	16	US-10-425-115-308664	Sequence 308664,	360	34.9	347	10	US-09-904-805-148	Sequence 148, App
288	38	34.9	322	15	US-10-369-493-10753	Sequence 10753, A	361	34.9	347	10	US-09-904-938A-148	Sequence 148, App
289	38	34.9	326	15	US-10-282-122A-51926	Sequence 51926, A	362	34.9	347	10	US-09-906-722A-148	Sequence 148, App
290	38	34.9	338	15	US-10-425-114-58760	Sequence 58760, A	363	34.9	347	10	US-09-908-576-148	Sequence 148, App
291	38	34.9	343	16	US-10-767-701-38027	Sequence 38027, A	364	34.9	347	14	US-10-227-884-114	Sequence 114, App
292	38	34.9	344	15	US-10-424-599-200984	Sequence 200984,	365	34.9	347	14	US-10-230-163-114	Sequence 114, App
293	38	34.9	347	9	US-09-909-320-148	Sequence 148, App	366	34.9	347	14	US-10-230-338-114	Sequence 114, App
294	38	34.9	347	9	US-09-909-088B-148	Sequence 148, App	367	34.9	347	14	US-10-218-631-114	Sequence 114, App
295	38	34.9	347	9	US-09-905-291A-148	Sequence 148, App	368	34.9	347	14	US-10-210-414-114	Sequence 114, App
296	38	34.9	347	9	US-09-902-853-148	Sequence 148, App	369	34.9	347	14	US-10-232-224-114	Sequence 114, App
297	38	34.9	347	9	US-09-907-841-148	Sequence 148, App	370	34.9	347	14	US-10-216-159A-114	Sequence 114, App
298	38	34.9	347	9	US-09-907-841-148	Sequence 148, App	371	34.9	347	14	US-10-218-849-114	Sequence 114, App
299	38	34.9	347	10	US-09-904-011-148	Sequence 148, App	372	34.9	347	14	US-10-227-873-114	Sequence 114, App
300	38	34.9	347	10	US-09-903-640-148	Sequence 148, App	373	34.9	347	14	US-10-227-883-114	Sequence 114, App
301	38	34.9	347	10	US-09-908-093-148	Sequence 148, App	374	34.9	347	14	US-10-219-076-114	Sequence 114, App
302	38	34.9	347	10	US-09-906-742-148	Sequence 148, App	375	34.9	347	14	US-10-230-434-114	Sequence 114, App
303	38	34.9	347	10	US-09-906-838-148	Sequence 148, App	376	34.9	347	14	US-10-219-003-114	Sequence 114, App

377	38	34.9	347	14	US-10-219-075-114	Sequence 114, App	450	38	34.9	347	14	US-10-216-163-114	Sequence 114, App
378	38	34.9	347	14	US-10-219-464-114	Sequence 114, App	451	38	34.9	347	14	US-10-299-976-148	Sequence 148, App
379	38	34.9	347	14	US-10-219-466-114	Sequence 114, App	452	38	34.9	347	14	US-10-239-937-148	Sequence 148, App
380	38	34.9	347	14	US-10-219-479-114	Sequence 114, App	453	38	34.9	347	14	US-10-223-081-24	Sequence 24, Appl
381	38	34.9	347	14	US-10-219-481-114	Sequence 114, App	454	38	34.9	347	14	US-10-218-765-114	Sequence 114, App
382	38	34.9	347	14	US-10-230-260-114	Sequence 114, App	455	38	34.9	347	14	US-10-219-063-114	Sequence 114, App
383	38	34.9	347	14	US-10-232-231-114	Sequence 114, App	456	38	34.9	347	14	US-10-219-066-114	Sequence 114, App
384	38	34.9	347	14	US-10-232-233-114	Sequence 114, App	457	38	34.9	347	14	US-10-219-067-114	Sequence 114, App
385	38	34.9	347	14	US-10-216-165-114	Sequence 114, App	458	38	34.9	347	14	US-10-219-068-114	Sequence 114, App
386	38	34.9	347	14	US-10-218-956-114	Sequence 114, App	459	38	34.9	347	14	US-10-219-069-114	Sequence 114, App
387	38	34.9	347	14	US-10-219-468-114	Sequence 114, App	460	38	34.9	347	14	US-10-219-073-114	Sequence 114, App
388	38	34.9	347	14	US-10-219-478-114	Sequence 114, App	461	38	34.9	347	14	US-10-219-475-114	Sequence 114, App
389	38	34.9	347	14	US-10-219-536-114	Sequence 114, App	462	38	34.9	347	14	US-10-219-480-114	Sequence 114, App
390	38	34.9	347	14	US-10-233-205-114	Sequence 114, App	463	38	34.9	347	14	US-10-219-483-114	Sequence 114, App
391	38	34.9	347	14	US-10-219-072-114	Sequence 114, App	464	38	34.9	347	14	US-10-219-525-114	Sequence 114, App
392	38	34.9	347	14	US-10-219-470-114	Sequence 114, App	465	38	34.9	347	14	US-10-219-526-114	Sequence 114, App
393	38	34.9	347	14	US-10-219-474-114	Sequence 114, App	466	38	34.9	347	14	US-10-219-530-114	Sequence 114, App
394	38	34.9	347	14	US-10-219-524-114	Sequence 114, App	467	38	34.9	347	14	US-10-219-531-114	Sequence 114, App
395	38	34.9	347	14	US-10-219-528-114	Sequence 114, App	468	38	34.9	347	14	US-10-219-532-114	Sequence 114, App
396	38	34.9	347	14	US-10-227-880-114	Sequence 114, App	469	38	34.9	347	14	US-10-219-533-114	Sequence 114, App
397	38	34.9	347	14	US-10-227-881-114	Sequence 114, App	470	38	34.9	347	14	US-10-230-437-114	Sequence 114, App
398	38	34.9	347	14	US-10-227-882-114	Sequence 114, App	471	38	34.9	347	14	US-10-232-228-114	Sequence 114, App
399	38	34.9	347	14	US-10-230-436-114	Sequence 114, App	472	38	34.9	347	14	US-10-232-082-24	Sequence 24, Appl
400	38	34.9	347	14	US-10-232-223-114	Sequence 114, App	473	38	34.9	347	15	US-10-298-993-148	Sequence 148, App
401	38	34.9	347	14	US-10-232-225-114	Sequence 114, App	474	38	34.9	347	15	US-10-305-654-24	Sequence 24, Appl
402	38	34.9	347	14	US-10-232-227-114	Sequence 114, App	475	38	34.9	347	15	US-10-448-923-148	Sequence 148, App
403	38	34.9	347	14	US-10-232-229-114	Sequence 114, App	476	38	34.9	347	15	US-10-449-656-148	Sequence 148, App
404	38	34.9	347	14	US-10-232-234-114	Sequence 114, App	477	38	34.9	347	15	US-10-232-226-114	Sequence 114, App
405	38	34.9	347	14	US-10-219-060-114	Sequence 114, App	478	38	34.9	347	15	US-10-448-713-148	Sequence 148, App
406	38	34.9	347	14	US-10-223-085-24	Sequence 24, Appl	479	38	34.9	347	15	US-10-230-130-114	Sequence 114, App
407	38	34.9	347	14	US-10-216-160-114	Sequence 114, App	480	38	34.9	347	15	US-10-425-447-148	Sequence 148, App
408	38	34.9	347	14	US-10-216-162-114	Sequence 114, App	481	38	34.9	347	15	US-10-081-056-24	Sequence 24, Appl
409	38	34.9	347	14	US-10-216-164-114	Sequence 114, App	482	38	34.9	347	15	US-10-219-535-114	Sequence 114, App
410	38	34.9	347	14	US-10-216-167-114	Sequence 114, App	483	38	34.9	347	15	US-10-232-230-114	Sequence 114, App
411	38	34.9	347	14	US-10-216-168-114	Sequence 114, App	484	38	34.9	347	15	US-10-119-480-114	Sequence 114, App
412	38	34.9	347	14	US-10-219-065-114	Sequence 114, App	485	38	34.9	347	16	US-10-215-371-148	Sequence 148, App
413	38	34.9	347	14	US-10-219-071-114	Sequence 114, App	486	38	34.9	347	16	US-10-771-187-148	Sequence 148, App
414	38	34.9	347	14	US-10-219-074-114	Sequence 114, App	487	38	34.9	347	16	US-10-219-477-114	Sequence 114, App
415	38	34.9	347	14	US-10-219-077-114	Sequence 114, App	488	38	34.9	347	17	US-10-963-467-148	Sequence 148, App
416	38	34.9	347	14	US-10-219-465-114	Sequence 114, App	489	38	34.9	347	17	US-10-978-255-148	Sequence 148, App
417	38	34.9	347	14	US-10-219-467-114	Sequence 114, App	490	38	34.9	347	18	US-10-797-366-148	Sequence 148, App
418	38	34.9	347	14	US-10-219-469-114	Sequence 114, App	491	38	34.9	355	15	US-10-425-114-49444	Sequence 49444, A
419	38	34.9	347	14	US-10-219-471-114	Sequence 114, App	492	38	34.9	367	16	US-10-739-930-7223	Sequence 7243, Ap
420	38	34.9	347	14	US-10-219-473-114	Sequence 114, App	493	38	34.9	389	17	US-10-491-823-322	Sequence 322, App
421	38	34.9	347	14	US-10-219-476-114	Sequence 114, App	494	38	34.9	391	11	US-09-939-853A-93	Sequence 93, Appl
422	38	34.9	347	14	US-10-219-482-114	Sequence 114, App	495	38	34.9	391	15	US-10-074-978A-402	Sequence 402, App
423	38	34.9	347	14	US-10-230-433-114	Sequence 114, App	496	38	34.9	394	15	US-10-425-114-42060	Sequence 42060, A
424	38	34.9	347	14	US-10-230-435-114	Sequence 114, App	497	38	34.9	407	15	US-10-282-122A-71177	Sequence 71177, A
425	38	34.9	347	14	US-10-230-438-114	Sequence 114, App	498	38	34.9	411	11	US-09-939-853A-92	Sequence 92, Appl
426	38	34.9	347	14	US-10-229-974-114	Sequence 114, App	499	38	34.9	411	15	US-10-074-978A-100	Sequence 100, App
427	38	34.9	347	14	US-10-230-024-114	Sequence 114, App	500	38	34.9	427	18	US-10-724-972A-6095	Sequence 6095, Ap
428	38	34.9	347	14	US-10-230-113-114	Sequence 114, App							
429	38	34.9	347	14	US-10-230-183-114	Sequence 114, App							
430	38	34.9	347	14	US-10-230-234-114	Sequence 114, App							
431	38	34.9	347	14	US-10-230-306-114	Sequence 114, App							
432	38	34.9	347	14	US-10-230-426-114	Sequence 114, App							
433	38	34.9	347	14	US-10-230-427-114	Sequence 114, App							
434	38	34.9	347	14	US-10-230-433-114	Sequence 114, App							
435	38	34.9	347	14	US-10-230-435-114	Sequence 114, App							
436	38	34.9	347	14	US-10-230-438-114	Sequence 114, App							
437	38	34.9	347	14	US-10-232-222-114	Sequence 114, App							
438	38	34.9	347	14	US-10-223-084-24	Sequence 24, Appl							
439	38	34.9	347	14	US-10-223-088-24	Sequence 24, Appl							
440	38	34.9	347	14	US-10-223-090-24	Sequence 24, Appl							
441	38	34.9	347	14	US-10-219-070-114	Sequence 114, App							
442	38	34.9	347	14	US-10-219-472-114	Sequence 114, App							
443	38	34.9	347	14	US-10-219-527-114	Sequence 114, App							
444	38	34.9	347	14	US-10-227-877-114	Sequence 114, App							
445	38	34.9	347	14	US-10-223-087-24	Sequence 24, Appl							
446	38	34.9	347	14	US-10-223-083-24	Sequence 24, Appl							
447	38	34.9	347	14	US-10-216-166-114	Sequence 114, App							
448	38	34.9	347	14	US-10-218-612-114	Sequence 114, App							
449	38	34.9	347	14	US-10-223-089-24	Sequence 24, Appl							

ALIGNMENTS

RESULT 1

US-10-812-238A-2

; Sequence 2, Application US/10812238A

; Publication No. US20050002904A1

; GENERAL INFORMATION:

; APPLICANT: Warty, Kishore, K.

; APPLICANT: Hunt-soe, Joseph O.

; TITLE OF INVENTION: Uses of Vascular Endothelial Growth Factor

; TITLE OF INVENTION: and Type I Collagen Inducible Protein (VCIP)

; FILE REFERENCE: D6563

; CURRENT APPLICATION NUMBER: US/10/812,238A

; CURRENT FILING DATE: 2004-03-29

; PRIOR APPLICATION NUMBER: US 60/458,164

; PRIOR FILING DATE: 2003-03-27

; NUMBER OF SEQ ID NOS: 36

; SEQ ID NO 2

; LENGTH: 20

; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: peptide used to raise anti-VCIP-RGD antibody
US-10-812-238A-2

Query Match 100.0%; Score 109; DB 17; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGYIYRCRGDDSKVQEAR 20
Db 1 EGYIYRCRGDDSKVQEAR 20

RESULT 2
US-10-655-601-2
; Sequence 2, Application US/10655601
; Publication No. US20040137522A1
; GENERAL INFORMATION:
; APPLICANT: Feany, Mel B.
; APPLICANT: Shulman, Joshua M.
; TITLE OF INVENTION: Genes and Proteins Altering Tau-Related Neuropathy
; FILE REFERENCE: 7570/73251
; CURRENT APPLICATION NUMBER: US/10/655,601
; CURRENT FILING DATE: 2003-09-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-655-601-2

Query Match 100.0%; Score 109; DB 16; Length 311;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGYIYRCRGDDSKVQEAR 20
Db 173 EGYIYRCRGDDSKVQEAR 192

RESULT 3
US-10-812-238A-13
; Sequence 13, Application US/10812238A
; Publication No. US20050002904A1
; GENERAL INFORMATION:
; APPLICANT: Wary, Kishore, K.
; APPLICANT: Humtsoe, Joseph O.
; TITLE OF INVENTION: Uses of Vascular Endothelial Growth Factor
; TITLE OF INVENTION: and Type I Collagen Inducible Protein (VCIP)
; FILE REFERENCE: D6563
; CURRENT APPLICATION NUMBER: US/10/812,238A
; CURRENT FILING DATE: 2004-03-29
; PRIOR APPLICATION NUMBER: US 60/458,164
; PRIOR FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 36
; SEQ ID NO 13
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: human VCIP
US-10-812-238A-13

Query Match 100.0%; Score 109; DB 17; Length 311;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGYIYRCRGDDSKVQEAR 20

Db 173 EGYIYRCRGDDSKVQEAR 192

RESULT 4
US-10-812-238A-23
; Sequence 23, Application US/10812238A
; Publication No. US20050002904A1
; GENERAL INFORMATION:
; APPLICANT: Wary, Kishore, K.
; APPLICANT: Humtsoe, Joseph O.
; TITLE OF INVENTION: Uses of Vascular Endothelial Growth Factor
; TITLE OF INVENTION: and Type I Collagen Inducible Protein (VCIP)
; FILE REFERENCE: D6563
; CURRENT APPLICATION NUMBER: US/10/812,238A
; CURRENT FILING DATE: 2004-03-29
; PRIOR APPLICATION NUMBER: US 60/458,164
; PRIOR FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 36
; SEQ ID NO 23
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: a peptide containing RGD sequence
US-10-812-238A-23

Query Match 67.0%; Score 73; DB 17; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00027;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 NYRCRGDDSKVQE 18
Db 1 NYRCRGDDSKVQE 13

RESULT 5
US-10-812-238A-20
; Sequence 20, Application US/10812238A
; Publication No. US20050002904A1
; GENERAL INFORMATION:
; APPLICANT: Wary, Kishore, K.
; APPLICANT: Humtsoe, Joseph O.
; TITLE OF INVENTION: Uses of Vascular Endothelial Growth Factor
; TITLE OF INVENTION: and Type I Collagen Inducible Protein (VCIP)
; FILE REFERENCE: D6563
; CURRENT APPLICATION NUMBER: US/10/812,238A
; CURRENT FILING DATE: 2004-03-29
; PRIOR APPLICATION NUMBER: US 60/458,164
; PRIOR FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 36
; SEQ ID NO 20
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: a peptide containing RGD sequence
US-10-812-238A-20

Query Match 54.1%; Score 59; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.029;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 NYRCRGDDSK 15
Db 1 NYRCRGDDSK 10

RESULT 6
US-10-287-226-346
; Sequence 346, Application US/10287226

```
; Publication No. US20040086875A1
; GENERAL INFORMATION:
; APPLICANT: Agee, Michele L.,
; APPLICANT: Alsobrook, John P.,
; APPLICANT: Berghs, Constance,
; APPLICANT: Boldog, Ference,
; APPLICANT: Burgess, Catherine E.,
; APPLICANT: Chant, John S.,
; APPLICANT: Chaudhuri, Amitabha,
; APPLICANT: DiPippo, Vincent A.,
; APPLICANT: Edinger, Shlomit R.,
; APPLICANT: Eisen, Andrew,
; APPLICANT: Ellerman, Karen,
; APPLICANT: Gangolli, Esha A.,
; APPLICANT: Gorman, Linda,
; APPLICANT: Gerlach, Valerie,
; APPLICANT: Ji, Weizhen,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Khrantsov, Nikolai,
; APPLICANT: Li, Li,
; APPLICANT: Malyankar, Uriel M.,
; APPLICANT: MacDougall, John R.,
; APPLICANT: Mezes, Peter S.,
; APPLICANT: Miller, Charles E.,
; APPLICANT: Millet, Isabelle,
; APPLICANT: Ooi, Chean Eng,
; APPLICANT: Ort, Tatiana,
; APPLICANT: Padigaru, Muralidhara,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Rastelli, Luca,
; APPLICANT: Rieger, Daniel K.,
; APPLICANT: Rothenberg, Mark E.,
; APPLICANT: Shenoy, Suresh G.,
; APPLICANT: Spaderna, Steven K.,
; APPLICANT: Spytek, Kimberley A.,
; APPLICANT: Taupier, Jr., Raymond J.,
; APPLICANT: Vernet, Corine A.M.,
; APPLICANT: Zerhusen, Bryan D.,
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-480C
; CURRENT APPLICATION NUMBER: US/10/287,226
; CURRENT FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: 60/334,421
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/354,392
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 60/360,148
; PRIOR FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: 60/364,000
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/404,821
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: 60/334,526
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/334,409
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/331,641
; PRIOR FILING DATE: 2001-11-20
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 673
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 346
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-287-226-346
; Query Match 54.1%; Score 59; DB 15; Length 221;
```

Best Local Similarity 50.0%; Pred. No. 0.68;
Matches 10; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 EGYIONVRCRGDDSKVQEAR 20
:||||:||||:|:|:|
DB 82 DGYIEYTCRGNAERVKEGR 101

RESULT 7
US-10-287-226-348
; Sequence 348, Application US/10287226
; Publication No. US20040086875A1
; GENERAL INFORMATION:
; APPLICANT: Agee, Michele L.,
; APPLICANT: Alsobrook, John P.,
; APPLICANT: Berghs, Constance,
; APPLICANT: Boldog, Ference,
; APPLICANT: Burgess, Catherine E.,
; APPLICANT: Chant, John S.,
; APPLICANT: Chaudhuri, Amitabha,
; APPLICANT: DiPippo, Vincent A.,
; APPLICANT: Edinger, Shlomit R.,
; APPLICANT: Eisen, Andrew,
; APPLICANT: Ellerman, Karen,
; APPLICANT: Gangolli, Esha A.,
; APPLICANT: Gorman, Linda,
; APPLICANT: Gerlach, Valerie,
; APPLICANT: Ji, Weizhen,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Khrantsov, Nikolai,
; APPLICANT: Li, Li,
; APPLICANT: Malyankar, Uriel M.,
; APPLICANT: MacDougall, John R.,
; APPLICANT: Mezes, Peter S.,
; APPLICANT: Miller, Charles E.,
; APPLICANT: Millet, Isabelle,
; APPLICANT: Ooi, Chean Eng,
; APPLICANT: Ort, Tatiana,
; APPLICANT: Padigaru, Muralidhara,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Rastelli, Luca,
; APPLICANT: Rieger, Daniel K.,
; APPLICANT: Rothenberg, Mark E.,
; APPLICANT: Shenoy, Suresh G.,
; APPLICANT: Spaderna, Steven K.,
; APPLICANT: Spytek, Kimberley A.,
; APPLICANT: Taupier, Jr., Raymond J.,
; APPLICANT: Vernet, Corine A.M.,
; APPLICANT: Zerhusen, Bryan D.,
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-480C
; CURRENT APPLICATION NUMBER: US/10/287,226
; CURRENT FILING DATE: 2002-11-04
; PRIOR APPLICATION NUMBER: 60/334,421
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/354,392
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 60/360,148
; PRIOR FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: 60/364,000
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/404,821
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: 60/334,526
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/334,409
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/331,641
; PRIOR FILING DATE: 2001-11-20
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 673
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 346
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-287-226-346

; PRIOR FILING DATE: 2001-11-20
; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 673

; SOFTWARE: CuraseqList version 0.1

; SEQ ID NO 348

; LENGTH: 221

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-287-226-348

Query Match 54.1%; Score 59; DB 15; Length 221;
Best Local Similarity 50.0%; Pred. No. 0.68; 5; Indels 0; Gaps 0;
Matches 10; Conservative 5; Mismatches 0; Indels 5; Gaps 0;

Qy 1 EGYIONRCRGDDSKVQEAR 20

:|||: |||: :|||:

Db 82 DGYIYYICRGNAERVKEGR 101

RESULT 8

US-10-764-425-157

; Sequence 157, Application US/10764425

; Publication No. US20040146921A1

; GENERAL INFORMATION:

; APPLICANT: Bayer Pharmaceuticals Corporation

; APPLICANT: Eveleigh, Deepa

; APPLICANT: Bigwood, Douglas

; APPLICANT: Taylor, Ian

; TITLE OF INVENTION: EXPRESSION PROFILES FOR COLON CANCER AND METHODS OF USE

; FILE REFERENCE: 5151

; CURRENT APPLICATION NUMBER: US/10764,425

; PRIOR FILING DATE: 2004-01-23

; PRIOR APPLICATION NUMBER: 60/442,582

; PRIOR FILING DATE: 2003-01-24

; NUMBER OF SEQ ID NOS: 191

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 157

; LENGTH: 284

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-764-425-157

Query Match 54.1%; Score 59; DB 18; Length 284;
Best Local Similarity 50.0%; Pred. No. 0.88;
Matches 10; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 EGYIONRCRGDDSKVQEAR 20

:|||: |||: :|||:

Db 145 DGYIYYICRGNAERVKEGR 164

RESULT 9

US-10-655-601-1

; Sequence 1, Application US/10655601

; Publication No. US20040137522A1

; GENERAL INFORMATION:

; APPLICANT: Feany, Mel B.

; APPLICANT: Shulman, Joshua M.

; TITLE OF INVENTION: Genes and Proteins Altering Tau-Related Neuropathy

; FILE REFERENCE: 7570/73251

; CURRENT APPLICATION NUMBER: US/10/655,601

; CURRENT FILING DATE: 2003-09-05

; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 285

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-655-601-1

Query Match 54.1%; Score 59; DB 16; Length 285;
Best Local Similarity 50.0%; Pred. No. 0.88;
Matches 10; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 EGYIONRCRGDDSKVQEAR 20

:|||: |||: :|||:

Db 146 DGYIYYICRGNAERVKEGR 165

RESULT 10

US-10-643-795A-136

; Sequence 136, Application US/10643795A

; Publication No. US20040241703A1

; GENERAL INFORMATION:

; APPLICANT: FREDERIC J. DESAUVAGE

; APPLICANT: GRETCHEN FRANTZ

; APPLICANT: KENNETH J. HILLAN

; APPLICANT: PAUL POLAKIS

; APPLICANT: ANDREW POLSON

; APPLICANT: VICTORIA SMITH

; APPLICANT: SUSAN D. SPENCER

; APPLICANT: THOMAS D. WU

; APPLICANT: ZEMIN ZHANG

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND

; FILE REFERENCE: P5026R1-US

; CURRENT APPLICATION NUMBER: US/10/643,795A

; PRIOR FILING DATE: 2003-08-19

; PRIOR APPLICATION NUMBER: US 60/404,809

; PRIOR FILING DATE: 2002-08-19

; PRIOR APPLICATION NUMBER: US 60/405,645

; PRIOR FILING DATE: 2002-08-21

; PRIOR APPLICATION NUMBER: US 60/413,192

; PRIOR FILING DATE: 2002-09-23

; PRIOR APPLICATION NUMBER: US 60/419,008

; PRIOR FILING DATE: 2002-10-15

; PRIOR APPLICATION NUMBER: US 60/426,847

; PRIOR FILING DATE: 2002-11-15

; PRIOR APPLICATION NUMBER: US 60/484,959

; PRIOR FILING DATE: 2003-07-02

; NUMBER OF SEQ ID NOS: 158

; SEQ ID NO 136

; LENGTH: 289

; TYPE: PRT

; ORGANISM: Homo sapien

US-10-643-795A-136

Query Match 54.1%; Score 59; DB 16; Length 289;
Best Local Similarity 50.0%; Pred. No. 0.89;
Matches 10; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 EGYIONRCRGDDSKVQEAR 20

:|||: |||: :|||:

Db 150 DGYIYYICRGNAERVKEGR 169

RESULT 11

US-10-948-518-136

; Sequence 136, Application US/10948518

; Publication No. US20050064492A1

; GENERAL INFORMATION:

; APPLICANT: FREDERIC J. DESAUVAGE

; APPLICANT: GRETCHEN FRANTZ

; APPLICANT: KENNETH J. HILLAN

; APPLICANT: PAUL POLAKIS

; APPLICANT: ANDREW POLSON

; APPLICANT: VICTORIA SMITH

; APPLICANT: SUSAN D. SPENCER

; APPLICANT: THOMAS D. WU

; APPLICANT: ZEMIN ZHANG

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND

; FILE REFERENCE: P5026R1-US

; CURRENT APPLICATION NUMBER: US/10/948,518

; CURRENT FILING DATE: 2004-09-22

; PRIOR APPLICATION NUMBER: US/10/643,795

```
; PRIOR FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: US 60/404,809
; PRIOR FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US 60/405,645
; PRIOR FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: US 60/413,192
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: US 60/419,008
; PRIOR FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/426,847
; PRIOR FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 60/484,959
; PRIOR FILING DATE: 2003-07-02
; NUMBER OF SEQ ID NOS: 158
; SEQ ID NO 136
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-948-518-136
```

```
Query Match 54.1%; Score 59; DB 17; Length 289;
Best Local Similarity 50.0%; Pred. No. 0.89;
Matches 10; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 EGYIQRRCRGDDSKVQEAR 20
      :|||: |||: |||: |||:
Db 150 DGVIEYICRGNAERVKEGR 169
```

```
RESULT 12
US-10-812-238A-22
; Sequence 22, Application US/10812238A
; Publication No. US20050002904A1
; GENERAL INFORMATION:
; APPLICANT: Wary, Kishore, K.
; APPLICANT: Humtsoe, Joseph O.
; TITLE OF INVENTION: Uses of Vascular Endothelial Growth Factor
; TITLE OF INVENTION: and Type I Collagen Inducible Protein (VCIP)
; FILE REFERENCE: D6563
; CURRENT APPLICATION NUMBER: US/10/812,238A
; CURRENT FILING DATE: 2004-03-29
; PRIOR APPLICATION NUMBER: US 60/458,164
; PRIOR FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 36
; SEQ ID NO 22
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: a peptide containing a mutated RGD sequence
US-10-812-238A-22
```

```
Query Match 50.5%; Score 55; DB 17; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.12;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 6 NYRCRGDDSK 15
      |||: |||: |||: |||:
Db 1 NYRCRGDDSK 10
```

```
RESULT 13
US-10-812-238A-21
; Sequence 21, Application US/10812238A
; Publication No. US20050002904A1
; GENERAL INFORMATION:
; APPLICANT: Wary, Kishore, K.
; APPLICANT: Humtsoe, Joseph O.
; TITLE OF INVENTION: Uses of Vascular Endothelial Growth Factor
; TITLE OF INVENTION: and Type I Collagen Inducible Protein (VCIP)
; FILE REFERENCE: D6563
; CURRENT APPLICATION NUMBER: US/10/812,238A
```

```
; CURRENT FILING DATE: 2004-03-29
; PRIOR APPLICATION NUMBER: US 60/458,164
; PRIOR FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 36
; SEQ ID NO 21
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: a peptide containing a mutated RGD sequence
US-10-812-238A-21
```

```
Query Match 48.6%; Score 53; DB 17; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.24;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 6 NYRCRGDDSK 15
      |||: |||: |||: |||:
Db 1 NYRCRADDSK 10
```

```
RESULT 14
US-10-369-493-13157
; Sequence 13157, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xiaofeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 13157
; LENGTH: 847
; TYPE: PRT
; ORGANISM: Aspergillus nidulans
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(847)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-13157
```

```
Query Match 45.0%; Score 49; DB 15; Length 847;
Best Local Similarity 53.3%; Pred. No. 92;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 2 GYIQNYRCRGDDSKV 16
      :|||: |||: |||: |||:
Db 181 GFQIKRCRGQDIEI 195
```

```
RESULT 15
US-10-369-493-13172
; Sequence 13172, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xiaofeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
```



```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: VAP amino acid
; OTHER INFORMATION: sequence of Imab11
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)..(135)
US-10-732-180-17

Query Match          41.3%; Score 45; DB 17; Length 135;
Best Local Similarity 53.3%; Pred. No. 58;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY      2 GYQNYRCRGDDSKV 16
      |||:|||||
Db      119 GYDSHYRCOGTDVTV 133

RESULT 20
US-09-147-947-6
; Sequence 6, Application US/09147947A
; Patent No. US20020160490A1
; GENERAL INFORMATION:
; APPLICANT: TSURUOKA, No. US20020160490A1uo
; APPLICANT: YAMASHIRO, Kyoto
; APPLICANT: YAMAGUCHI, No. US20020160490A1om1
; TITLE OF INVENTION: No. US20020160490A1el Serine Protease
; FILE REFERENCE: 001560-349
; CURRENT APPLICATION NUMBER: US/09/147,947A
; CURRENT FILING DATE: 1997-03-24
; EARLIER APPLICATION NUMBER: PCT/JP98/03324
; EARLIER FILING DATE: 1998-07-24
; EARLIER APPLICATION NUMBER: JP 9/213969
; EARLIER FILING DATE: 1997-07-24
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 6
; LENGTH: 822
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; OTHER INFORMATION:
US-09-147-947-6

Query Match          41.3%; Score 45; DB 9; Length 822;
Best Local Similarity 50.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY      3 YQNYRCRGDDSKV 16
      |||:|||||
Db      180 YWSNVRRCGDEENI 193

RESULT 21
US-10-757-262-34
; Sequence 34, Application US/10757262
; Publication No. US20040197825A1
; GENERAL INFORMATION:
; APPLICANT: Karicheti, Venkateswarlu
; APPLICANT: Silos-Santiago, Immaculada
; APPLICANT: Eliaaof, Scott D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: UROLOGICAL DISORDERS USING 44390, 54181, 211, 5687, 884,
; TITLE OF INVENTION: 1405, 636, 4421, 5410, 30905, 2045, 16405, 18560, 2047,
; TITLE OF INVENTION: 33751, 52872, 14063, 20739, 32544, 43239, 44373, 51164,
; TITLE OF INVENTION: 53010, 16852, 1587, 2207, 2245, 2387, 52908, 69112, 14990,
; TITLE OF INVENTION: 32720, 115, 579, 15985, 15625, 760, 18603, 2395, 2554, 8675,
; TITLE OF INVENTION: 32720, 4809, 14303, 16816, 17827, 32620, 577, 619, 1423,
; TITLE OF INVENTION: 2158, 8263, 15402, 16209, 16386, 21165, 30911, 41897, 1643,
; TITLE OF INVENTION: 2543, 9626, 13231, 32409, 84260, 2882, 8203, 32678 OR
; TITLE OF INVENTION: 55053
; FILE REFERENCE: MPI03-007PIRNOVMNIM
; CURRENT APPLICATION NUMBER: US/10/757,262

; CURRENT FILING DATE: 2004-01-14
; PRIOR APPLICATION NUMBER: US 60/440,318
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: US 60/444,783
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: US 60/457,901
; PRIOR FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: US 60/468,775
; PRIOR FILING DATE: 2003-05-08
; PRIOR APPLICATION NUMBER: US 60/471,614
; PRIOR FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: US 60/478,742
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/488,529
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/491,156
; PRIOR FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: US 60/499,594
; PRIOR FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 60/506,332
; PRIOR FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 136
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 875
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-757-262-34

Query Match          41.3%; Score 45; DB 16; Length 875;
Best Local Similarity 50.0%; Pred. No. 3.9e+02;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY      3 YQNYRCRGDDSKV 16
      |||:|||||
Db      233 YWSNVRRCGDEENI 246

RESULT 22
US-10-723-860-4046
; Sequence 4046, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 4046
; LENGTH: 875
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-723-860-4046

Query Match          41.3%; Score 45; DB 16; Length 875;
Best Local Similarity 50.0%; Pred. No. 3.9e+02;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY      3 YQNYRCRGDDSKV 16
      |||:|||||
Db      233 YWSNVRRCGDEENI 246

RESULT 23
US-10-843-299-2
; Sequence 2, Application US/10843299
```

APPLICANT: Kekuda, Ramesh

RESULT 26
US-10-108-260A-2858
; Sequence 2858, Application US/10108260A
; Publication No. US20040005560A1

```

RESULT 25
US-10-114-270-104
; Sequence 104, Application US/10114270
; Publication No. US20040030110A1
; GENERAL INFORMATION:
; APPLICANT: Guo, Xiaojia
; APPLICANT: Kekuda, Ramesh

```

; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560a1e1 full length cdna
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2858
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-2858

Query Match 40.4%; Score 44; DB 15; Length 485;
Best Local Similarity 53.8%; Pred. No. 3.1e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 YIQNYRCRGDDSK 15
|:|:|:|:|:
Db 466 YLYNWRCPGTDSE 478

RESULT 27
US-10-427-631-1
; Sequence 1, Application US/10427631
; Publication No. US20030175923A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; TANG, Y. Tom;
; APPLICANT: CORLEY, Neil C.; GUEGLER, Karl J.;
; APPLICANT: BAUGHN, Mariah R.; LAL, Preeti G.;
; APPLICANT: YUE, Henry; HILLMAN, Jennifer L.;
; APPLICANT: AZIMZAI, Yalda
; TITLE OF INVENTION: HUMAN TRANSFERASE PROTEINS
; FILE REFERENCE: PP-0592-1 DIV
; CURRENT APPLICATION NUMBER: US/10/427,631
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: US 09/786,240
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: PCT/US99/20989
; PRIOR FILING DATE: 1999-09-09
; PRIOR APPLICATION NUMBER: US 60/172,220
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: US 60/155,248
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: US 60/133,642
; PRIOR FILING DATE: 1999-05-11
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030175923A1 1632930CD1
US-10-427-631-1

Query Match 40.4%; Score 44; DB 14; Length 498;
Best Local Similarity 53.8%; Pred. No. 3.1e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 YIQNYRCRGDDSK 15
|:|:|:|:|:
Db 479 YLYNWRCPGTDSE 491

RESULT 28
US-10-369-493-13176
; Sequence 13176, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei

; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 13176
; LENGTH: 1100
; TYPE: PRT
; ORGANISM: Aspergillus nidulans
US-10-369-493-13176

Query Match 40.4%; Score 44; DB 15; Length 1100;
Best Local Similarity 46.7%; Pred. No. 7.1e+02;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 GYIQNYRCRGDDSKV 16
|:|:|:|:|:
Db 59 GFIOQRKCHGQDIEI 73

RESULT 29
US-09-832-292-39
; Sequence 39, Application US/09832292
; Patent No. US20020177205A1
; GENERAL INFORMATION:
; APPLICANT: Ryazanov, Alexey
; TITLE OF INVENTION: MAMMALIAN ALPHA-KINASE PROTEINS, NUCLEIC ACIDS AND
; FILE REFERENCE: 601-1-098CIP
; CURRENT APPLICATION NUMBER: US/09/832,292
; CURRENT FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 09/632,131
; PRIOR FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39
; LENGTH: 1907
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-832-292-39

Query Match 40.4%; Score 44; DB 9; Length 1907;
Best Local Similarity 43.8%; Pred. No. 1.2e+03;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 4 IQNYRCRGDDSKVQEA 19
|:|:|:|:|:
Db 339 LQLYRCREDAAIYQA 354

RESULT 30
US-10-437-963-168323
; Sequence 168323, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B

```
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 168323
; LENGTH: 3993
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_66849C.1.pep
; US-10-437-963-168323

Query Match          40.4%; Score 44; DB 16; Length 3993;
Best Local Similarity 42.9%; Pred. No. 2.6e+03;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 4 IQNYRCRGDSDSKVQ 17
   :||:| |||:::
Db 132 VQNFPCASDDSEIK 145

RESULT 31
US-10-424-599-202541
; Sequence 202541, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 202541
; LENGTH: 130
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(130)
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: PAT_MRT3847_2491C.1.pep
; US-10-424-599-202541

Query Match          39.9%; Score 43.5; DB 15; Length 130;
Best Local Similarity 56.2%; Pred. No. 96;
Matches 9; Conservative 4; Mismatches 2; Indels 1; Gaps 1;

QY 3 YIQNYRCRG-DDSKVQ 17
   ||: : |||: |||
Db 91 YIKFPACRGSSNSKVQ 106

RESULT 32
US-10-767-701-36687
; Sequence 36687, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 36687
; LENGTH: 193
; TYPE: PRT

; ORGANISM: Sorghum bicolor
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(193)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C10059_1.pep
; US-10-767-701-36687

Query Match          39.4%; Score 43; DB 16; Length 193;
Best Local Similarity 53.8%; Pred. No. 1.7e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 GYIQNYRCRGDSDS 14
   ||: || |||:
Db 89 GYVTNYAVDGDGA 101

RESULT 33
US-10-291-172-610
; Sequence 610, Application US/10291172
; Publication No. US20030228584A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: No. US20030228584A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-045
; CURRENT APPLICATION NUMBER: US/10/291,172
; CURRENT FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 09/693,267
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,847
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/596,193
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 610
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-291-172-610

Query Match          39.4%; Score 43; DB 15; Length 237;
Best Local Similarity 58.3%; Pred. No. 2.1e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 IQNYRCRGDSDSK 15
   :||: ||| |||:
Db 3 LDNHRCRGDLTK 14

RESULT 34
US-10-221-278-610
; Sequence 610, Application US/10221278
; Publication No. US20040034208A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: No. US20040034208A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-045
; CURRENT APPLICATION NUMBER: US/10/221,278
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: 09/693,267
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,847
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/596,193
```

```
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 610
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-221-278-610
```

```
Query Match 39.4%; Score 43; DB 15; Length 237;
Best Local Similarity 58.3%; Pred. No. 2.1e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 4 IQNYRCRGDDSK 15
:|:|||||:|
Db 3 LDNHRRCGDLTK 14
```

```
RESULT 35
US-10-156-761-13203
; Sequence 13203, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 13203
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-13203
```

```
Query Match 39.4%; Score 43; DB 14; Length 363;
Best Local Similarity 46.7%; Pred. No. 3.2e+02;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 3 YIQNYRCRGDDSKVQ 17
|:|:| ||| |:|
Db 182 YLRGRTTGDGTGE 196
```

```
RESULT 36
US-10-210-281-12
; Sequence 12, Application US/10210281
; Publication No. US20040030096A1
; GENERAL INFORMATION:
; APPLICANT: Gorman, Linda
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Guo, Xiaojia
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Zhong, Mei
; APPLICANT: Patturajan, Meera
; APPLICANT: Miller, Charles E.
; APPLICANT: Ji, Weizhen
; APPLICANT: Pena, Carol E.A.
```

```
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Sciore, Paul
; APPLICANT: Stone, David J.
; APPLICANT: Taupier, Raymond J., Jr.
; APPLICANT: Casman, Stacie
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Boldog, Ferenc L.
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
; TITLE OF INVENTION: THE SAME
; FILE REFERENCE: 21402-416D
; CURRENT APPLICATION NUMBER: US/10/210,281
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: 60/309,501
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/310,291
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/361,775
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 60/310,951
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/361,832
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 60/311,292
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/311,979
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 60/312,203
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: 60/313,201
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/313,702
; PRIOR FILING DATE: 2001-08-20
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 12
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-210-281-12
```

```
Query Match 39.4%; Score 43; DB 15; Length 421;
Best Local Similarity 58.3%; Pred. No. 3.8e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 4 IQNYRCRGDDSK 15
:|:|||||:|
Db 195 LDNHRRCGDLTK 206
```

```
RESULT 37
US-10-210-281-10
; Sequence 10, Application US/10210281
; Publication No. US20040030096A1
; GENERAL INFORMATION:
; APPLICANT: Gorman, Linda
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Guo, Xiaojia
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Zhong, Mei
; APPLICANT: Patturajan, Meera
; APPLICANT: Miller, Charles E.
; APPLICANT: Ji, Weizhen
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Sciore, Paul
; APPLICANT: Stone, David J.
; APPLICANT: Taupier, Raymond J., Jr.
; APPLICANT: Casman, Stacie
; APPLICANT: Rothenberg, Mark E.
```



```
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Boldog, Ferenc L.
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS
; TITLE OF INVENTION: THE SAME
; FILE REFERENCE: 21402-416D
; CURRENT APPLICATION NUMBER: US/10/210,281
; PRIOR FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: 60/309,501
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/310,291
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/361,775
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 60/310,951
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/361,832
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 60/311,292
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/311,979
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 60/312,203
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: 60/313,201
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/313,702
; PRIOR FILING DATE: 2001-08-20
; REMAINING Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 10
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-210-281-10

Query Match      39.4%; Score 43; DB 15; Length 424;
Best Local Similarity 58.3%; Pred. No. 3.8e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      4 IONVRCRGDSDK 15
       :|:|||||:|
Db      196 LDNHRRCGDLTK 207

RESULT 38
US-10-157-031-139
; Sequence 139, Application US/10157031
; Publication No. US20030108890A1
; GENERAL INFORMATION:
; APPLICANT: Baranova, A. V.
; APPLICANT: Yankovsky, N. K.
; APPLICANT: Kozlov, A. P.
; APPLICANT: Lobahev, A. V.
; APPLICANT: Krukovskaya, L. L.
; TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequences
; FILE REFERENCE: 2760-103
; CURRENT APPLICATION NUMBER: US/10/157,031
; CURRENT FILING DATE: 2002-05-30
; NUMBER OF SEQ ID NOS: 415
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 139
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-157-031-139

Query Match      39.4%; Score 43; DB 14; Length 428;
Best Local Similarity 58.3%; Pred. No. 3.8e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      4 IONVRCRGDSDK 15
       :|:|||||:|
```

```
Db      195 LDNHRRCGDLTK 206

RESULT 39
US-10-108-260A-2988
; Sequence 2988, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2988
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-2988

Query Match      39.4%; Score 43; DB 15; Length 489;
Best Local Similarity 58.3%; Pred. No. 4.4e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      4 IONVRCRGDSDK 15
       :|:|||||:|
Db      256 LDNHRRCGDLTK 267

RESULT 40
US-10-369-493-21996
; Sequence 21996, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 21996
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-369-493-21996

Query Match      39.4%; Score 43; DB 15; Length 505;
Best Local Similarity 72.7%; Pred. No. 4.5e+02;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      10 RGDSDKVQEAR 20
       :|:|||||:|
Db      487 RSDAKVEEAR 497

RESULT 41
US-09-815-242-12804
; Sequence 12804, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
```

RESULT 44
US-10-282-287-4
; Sequence 4, Application US/10282287
; Publication No. US20030129633A1
; GENERAL INFORMATION:
; APPLICANT: O'Donnell, Michael E.
; APPLICANT: Zhang, Dan
; APPLICANT: Whipple, Richard
; TITLE OF INVENTION: DNA REPLICATION PROTEINS OF GRAM POSITIVE BACTERIA AND
; TITLE OF INVENTION: THEIR USE TO SCREEN FOR CHEMICAL INHIBITORS
; FILE REFERENCE: 22221/1002
; CURRENT APPLICATION NUMBER: US/10/282,287
; CURRENT FILING DATE: 2002-10-28
; PRIOR APPLICATION NUMBER: US/09/235,245
; PRIOR FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: 60/074,522
; PRIOR FILING DATE: 1998-01-27
; PRIOR APPLICATION NUMBER: 60/093,727


```
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 135805
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(208)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_37445C.1.pap
US-10-437-963-135805

Query Match      38.5%; Score 42; DB 16; Length 208;
Best Local Similarity 43.8%; Pred. No. 2.6e+02;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY      3 YIQNYRCRGDDSKVQE 18
Db      93 FMFDYSCGGDDSGTEE 108

RESULT 49
US-10-898-775-28
; Sequence 28, Application US/10898775
; Publication No. US20050142647A1
; GENERAL INFORMATION:
; APPLICANT: WASSELL, PAUL
; APPLICANT: SOE, JORN BORCH
; APPLICANT: MIKKELSON, JORN DALGAARD
; APPLICANT: KRISTENSEN, ANNA CECILIE JENTOFT
; TITLE OF INVENTION: METHOD
; FILE REFERENCE: 674509-2058
; CURRENT APPLICATION NUMBER: US/10/898,775
; CURRENT FILING DATE: 2004-07-26
; PRIOR APPLICATION NUMBER: GB 0416023.0
; PRIOR FILING DATE: 2004-07-16
; PRIOR APPLICATION NUMBER: PCT/IB04/00655
; PRIOR FILING DATE: 2004-01-15
; PRIOR APPLICATION NUMBER: GB 0330016.7
; PRIOR FILING DATE: 2003-12-24
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 28
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Streptomyces coelicolor
US-10-898-775-28

Query Match      38.5%; Score 42; DB 18; Length 305;
Best Local Similarity 61.5%; Pred. No. 3.9e+02;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      8 RCRGDDSKVQEAR 20
Db      258 RCRDDGAVHEPR 270

RESULT 50
US-10-437-963-135808
; Sequence 135808, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 135808
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(473)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_37448C.1.pap
US-10-437-963-135808

Query Match      38.5%; Score 42; DB 16; Length 473;
Best Local Similarity 43.8%; Pred. No. 6.1e+02;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY      3 YIQNYRCRGDDSKVQE 18
Db      150 FMFDYSCGGDDSGTEE 165

Search completed: September 7, 2005, 20:02:58
Job time : 143.558 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2005, 19:44:24 ; Search time 25.1163 Seconds
(without alignments)
59.443 Million cell updates/sec

Title: US-10-812-238B-2

Perfect score: 109

Sequence: 1 EGYIYRCRGDDSKVQEAR 20

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database :

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	109	100.0	232	4	US-09-360-376-55
2	66	60.6	282	4	US-09-360-376-54
3	65	59.6	283	3	US-08-992-035A-3
4	65	59.6	283	4	US-09-360-376-53
5	59	54.1	285	3	US-08-992-035A-1
6	46.5	42.7	705	4	US-08-311-731A-4
7	44	40.4	416	4	US-09-949-016-8237
8	44	40.4	498	4	US-09-786-240-1
9	43	39.4	505	4	US-09-538-092-359
10	42	38.5	167	4	US-09-252-991A-29214
11	42	38.5	132	4	US-09-248-796A-18072
12	41	37.6	132	4	US-09-543-681A-7222
13	41	37.6	201	3	US-09-015-734-12
14	41	37.6	201	4	US-09-015-311-12
15	41	37.6	236	3	US-09-015-734-7
16	41	37.6	236	4	US-09-015-311-7
17	41	37.6	255	3	US-09-015-734-2
18	41	37.6	255	4	US-09-515-311-2
19	41	37.6	272	4	US-09-489-039A-13748
20	41	37.6	396	4	US-09-506-286B-92
21	41	37.6	396	4	US-09-506-286B-95
22	41	37.6	396	4	US-10-065-133A-99
23	41	37.6	396	4	US-10-065-133A-95
24	41	37.6	757	4	US-09-506-286B-107
25	41	37.6	757	4	US-10-065-133A-107
26	40.5	37.2	145	4	US-09-002-540-13372
27	40.5	37.2	314	4	US-09-252-991A-20336

28	40	36.7	94	4	US-09-198-452A-1290	Sequence 1290, Ap
29	40	36.7	256	3	US-09-134-001C-5278	Sequence 5278, Ap
30	40	36.7	396	4	US-09-506-286B-86	Sequence 86, Appl
31	40	36.7	396	4	US-09-506-286B-89	Sequence 89, Appl
32	40	36.7	396	4	US-10-065-133A-86	Sequence 86, Appl
33	40	36.7	396	4	US-10-065-133A-89	Sequence 89, Appl
34	40	36.7	424	4	US-09-489-039A-12030	Sequence 12030, A
35	40	36.7	501	4	US-09-270-767-41177	Sequence 41177, A
36	40	36.7	710	4	US-09-270-767-56393	Sequence 56393, A
37	40	36.7	710	4	US-09-252-991A-19287	Sequence 19287, A
38	40	36.7	730	1	US-08-121-713D-58	Sequence 58, Appl
39	40	36.7	730	1	US-08-835-268-58	Sequence 58, Appl
40	40	36.7	730	2	US-09-060-692-58	Sequence 58, Appl
41	40	36.7	730	3	US-08-833-391-58	Sequence 58, Appl
42	40	36.7	730	3	US-09-060-610-58	Sequence 58, Appl
43	40	36.7	730	5	PCT-US94-10151A-58	Sequence 58, Appl
44	40	36.7	757	4	US-09-506-286B-104	Sequence 104, Appl
45	40	36.7	757	4	US-10-065-133A-104	Sequence 104, Appl
46	40	36.7	980	4	US-09-252-991A-30838	Sequence 30838, A
47	39.5	36.2	99	4	US-09-621-976-4220	Sequence 4220, Ap
48	39.5	36.2	137	3	US-09-188-930-174	Sequence 174, Appl
49	39.5	36.2	137	4	US-09-312-283C-174	Sequence 174, Appl
50	39.5	36.2	225	4	US-09-673-395A-354	Sequence 354, Appl
51	39.5	36.2	413	4	US-09-270-767-44418	Sequence 44418, A
52	39.5	36.2	456	3	US-09-134-001C-3659	Sequence 3659, Ap
53	39.5	36.2	1238	4	US-09-252-991A-26363	Sequence 26363, A
54	39	35.8	106	4	US-09-673-395A-263	Sequence 263, Appl
55	39	35.8	169	4	US-09-252-991A-19373	Sequence 19373, A
56	39	35.8	317	4	US-09-270-767-42366	Sequence 42366, A
57	39	35.8	447	3	US-09-627-376-10	Sequence 10, Appl
58	39	35.8	447	4	US-10-047-678B-10	Sequence 10, Appl
59	39	35.8	621	2	US-08-419-652-4	Sequence 4, Appl
60	39	35.8	660	1	US-08-248-532-3	Sequence 3, Appl
61	39	35.8	660	2	US-08-419-652-3	Sequence 3, Appl
62	39	35.8	662	1	US-08-248-532-2	Sequence 2, Appl
63	39	35.8	662	2	US-08-419-652-2	Sequence 2, Appl
64	39	35.8	662	2	US-08-685-118-4	Sequence 4, Appl
65	39	35.8	662	2	US-08-915-495-4	Sequence 4, Appl
66	39	35.8	662	2	US-08-914-520-4	Sequence 4, Appl
67	39	35.8	662	3	US-08-789-350-2	Sequence 2, Appl
68	39	35.8	662	4	US-09-949-016-6624	Sequence 6624, Ap
69	39	35.8	683	4	US-09-949-016-10092	Sequence 10092, A
70	39	35.8	917	1	US-08-245-295-2	Sequence 2, Appl
71	39	35.8	917	1	US-08-481-130-2	Sequence 2, Appl
72	39	35.8	917	1	US-08-656-984A-2	Sequence 2, Appl
73	39	35.8	917	1	US-08-485-604-2	Sequence 2, Appl
74	39	35.8	917	2	US-08-487-595-2	Sequence 2, Appl
75	39	35.8	924	1	US-08-481-130-28	Sequence 28, Appl
76	39	35.8	924	1	US-08-656-984A-28	Sequence 28, Appl
77	39	35.8	924	1	US-08-485-604-28	Sequence 28, Appl
78	39	35.8	924	2	US-08-487-595-28	Sequence 28, Appl
79	38.5	35.3	383	4	US-09-248-796A-18954	Sequence 18954, A
80	38	34.9	102	3	US-09-034-916-5	Sequence 5, Appl
81	38	34.9	120	4	US-09-513-999C-6407	Sequence 6407, Ap
82	38	34.9	136	4	US-09-690-454-131	Sequence 131, Appl
83	38	34.9	213	4	US-09-252-991A-22600	Sequence 22600, A
84	38	34.9	265	4	US-09-583-110-3445	Sequence 3445, Ap
85	38	34.9	284	4	US-09-134-000C-3594	Sequence 3594, Ap
86	38	34.9	346	3	US-09-034-916-2	Sequence 2, Appl
87	38	34.9	347	4	US-09-907-794A-148	Sequence 148, Appl
88	38	34.9	347	4	US-09-905-125A-148	Sequence 148, Appl
89	38	34.9	347	4	US-09-902-775A-148	Sequence 148, Appl
90	38	34.9	347	4	US-09-906-700-148	Sequence 148, Appl
91	38	34.9	347	4	US-09-903-603A-148	Sequence 148, Appl
92	38	34.9	347	4	US-09-904-920A-148	Sequence 148, Appl
93	38	34.9	347	4	US-09-909-064-148	Sequence 148, Appl
94	38	34.9	347	4	US-09-905-381A-148	Sequence 148, Appl
95	38	34.9	347	4	US-09-906-618-148	Sequence 148, Appl
96	38	34.9	369	4	US-09-949-016-10311	Sequence 10311, A
97	38	34.9	380	4	US-09-252-991A-28929	Sequence 28929, A
98	38	34.9	427	3	US-09-134-001C-5143	Sequence 5143, Ap
99	38	34.9	440	4	US-09-270-767-60249	Sequence 60249, A
100	38	34.9	456	1	US-08-464-164-2	Sequence 2, Appl

101	38	34.9	456	1	US-08-338-057-2	Sequence 2, Appli	174	36	33.0	35	3	US-09-082-279B-144	Sequence 144, App
102	38	34.9	456	2	US-08-668-416-2	Sequence 2, Appli	175	36	33.0	35	3	US-09-082-279B-145	Sequence 145, App
103	38	34.9	463	3	US-09-252-991A-31367	Sequence 31367, A	176	36	33.0	35	3	US-09-082-279B-146	Sequence 146, App
104	38	34.9	503	4	US-09-270-767-37684	Sequence 37684, A	177	36	33.0	35	3	US-09-082-279B-147	Sequence 147, App
105	38	34.9	503	4	US-09-270-767-52901	Sequence 52901, A	178	36	33.0	35	3	US-09-082-279B-148	Sequence 148, App
106	38	34.9	503	4	US-09-248-796A-37110	Sequence 17110, A	179	36	33.0	35	3	US-09-082-279B-149	Sequence 149, App
107	38	34.9	532	4	US-09-252-991A-21005	Sequence 21005, A	180	36	33.0	35	3	US-08-474-349A-228	Sequence 228, App
108	38	34.9	572	5	PC7-US91-08177-11	Sequence 11, Appl	181	36	33.0	35	3	US-09-315-304B-22	Sequence 22, Appl
109	38	34.9	625	4	US-09-252-991A-28537	Sequence 28537, A	182	36	33.0	35	3	US-09-315-304B-138	Sequence 138, App
110	38	34.9	923	4	US-09-717-926-2	Sequence 2, Appli	183	36	33.0	35	3	US-09-315-304B-139	Sequence 139, App
111	37.5	34.4	304	4	US-09-252-991A-26325	Sequence 26325, A	184	36	33.0	35	3	US-09-315-304B-140	Sequence 140, App
112	37.5	34.4	511	4	US-09-252-991A-26396	Sequence 26396, A	185	36	33.0	35	3	US-09-315-304B-141	Sequence 141, App
113	37.5	34.4	1269	3	US-08-781-891-74	Sequence 74, Appl	186	36	33.0	35	3	US-09-315-304B-142	Sequence 142, App
114	37.5	34.4	1269	4	US-09-618-166-74	Sequence 74, Appl	187	36	33.0	35	3	US-09-315-304B-143	Sequence 143, App
115	37.5	34.4	1409	4	US-09-949-001-22	Sequence 22, Appl	188	36	33.0	35	3	US-09-315-304B-144	Sequence 144, App
116	37.5	34.4	1432	3	US-08-781-891-71	Sequence 71, Appl	189	36	33.0	35	3	US-09-315-304B-145	Sequence 145, App
117	37.5	34.4	1432	4	US-09-618-166-71	Sequence 71, Appl	190	36	33.0	35	3	US-09-315-304B-146	Sequence 146, App
118	37.5	34.4	1432	4	US-09-949-001-16	Sequence 16, Appl	191	36	33.0	35	3	US-09-315-304B-147	Sequence 147, App
119	37	33.9	129	4	US-09-732-210-1649	Sequence 320, App	192	36	33.0	35	3	US-09-315-304B-148	Sequence 148, App
120	37	33.9	130	4	US-09-711-164-320	Sequence 318, App	193	36	33.0	35	3	US-09-315-304B-149	Sequence 149, App
121	37	33.9	132	4	US-09-492-709A-318	Sequence 11034, A	194	36	33.0	35	4	US-09-834-784-22	Sequence 22, Appl
122	37	33.9	132	4	US-09-489-039A-11034	Sequence 11034, A	195	36	33.0	35	4	US-09-834-784-138	Sequence 138, App
123	37	33.9	159	4	US-09-252-991A-16672	Sequence 16672, A	196	36	33.0	35	4	US-09-834-784-139	Sequence 139, App
124	37	33.9	185	4	US-09-328-352-7477	Sequence 7477, Ap	197	36	33.0	35	4	US-09-834-784-140	Sequence 140, App
125	37	33.9	191	4	US-09-602-777A-350	Sequence 350, App	198	36	33.0	35	4	US-09-834-784-141	Sequence 141, App
126	37	33.9	236	4	US-09-479-040-25	Sequence 25, Appl	199	36	33.0	35	4	US-09-834-784-142	Sequence 142, App
127	37	33.9	275	4	US-09-489-039A-11789	Sequence 11789, A	200	36	33.0	35	4	US-09-834-784-143	Sequence 143, App
128	37	33.9	332	4	US-09-585-858-50	Sequence 50, Appl	201	36	33.0	35	4	US-09-834-784-144	Sequence 144, App
129	37	33.9	332	4	US-10-270-878-50	Sequence 50, Appl	202	36	33.0	35	4	US-09-834-784-145	Sequence 145, App
130	37	33.9	341	4	US-09-252-991A-16890	Sequence 16890, A	203	36	33.0	35	4	US-09-834-784-146	Sequence 146, App
131	37	33.9	357	4	US-09-248-796A-21669	Sequence 21669, A	204	36	33.0	35	4	US-09-834-784-147	Sequence 147, App
132	37	33.9	362	4	US-09-252-991A-31016	Sequence 31016, A	205	36	33.0	35	4	US-09-834-784-148	Sequence 148, App
133	37	33.9	391	4	US-09-252-991A-16725	Sequence 16725, A	206	36	33.0	35	4	US-09-834-784-149	Sequence 149, App
134	37	33.9	400	1	US-08-447-500-6	Sequence 6, Appli	207	36	33.0	35	4	US-09-515-965A-22	Sequence 22, Appl
135	37	33.9	400	1	US-08-454-097-6	Sequence 6, Appli	208	36	33.0	35	4	US-09-515-965A-138	Sequence 138, App
136	37	33.9	400	1	US-08-453-866-6	Sequence 6, Appli	209	36	33.0	35	4	US-09-515-965A-139	Sequence 139, App
137	37	33.9	400	3	US-08-185-359-6	Sequence 6, Appli	210	36	33.0	35	4	US-09-515-965A-140	Sequence 140, App
138	37	33.9	403	4	US-09-252-991A-25048	Sequence 25048, A	211	36	33.0	35	4	US-09-515-965A-141	Sequence 141, App
139	37	33.9	437	4	US-09-262-537-8	Sequence 2, Appli	212	36	33.0	35	4	US-09-515-965A-142	Sequence 142, App
140	37	33.9	464	3	US-09-056-783-2	Sequence 2, Appli	213	36	33.0	35	4	US-09-515-965A-143	Sequence 143, App
141	37	33.9	488	4	US-09-489-039A-13164	Sequence 13164, A	214	36	33.0	35	4	US-09-515-965A-144	Sequence 144, App
142	37	33.9	583	4	US-09-270-767-43661	Sequence 43661, A	215	36	33.0	35	4	US-09-515-965A-145	Sequence 145, App
143	37	33.9	754	3	US-09-276-400-8	Sequence 8, Appli	216	36	33.0	35	4	US-09-515-965A-146	Sequence 146, App
144	37	33.9	754	3	US-09-448-076-8	Sequence 8, Appli	217	36	33.0	35	4	US-09-515-965A-147	Sequence 147, App
145	37	33.9	754	3	US-09-702-572-8	Sequence 8, Appli	218	36	33.0	35	4	US-09-515-965A-148	Sequence 148, App
146	37	33.9	944	4	US-09-270-767-46843	Sequence 46843, A	219	36	33.0	35	4	US-09-515-965A-149	Sequence 149, App
147	37	33.9	946	4	US-09-502-540-16817	Sequence 16817, A	220	36	33.0	35	4	US-09-350-641C-22	Sequence 22, Appl
148	37	33.9	1123	4	US-09-262-537-4	Sequence 4, Appli	221	36	33.0	35	4	US-09-350-641C-138	Sequence 138, App
149	37	33.9	1177	4	US-09-262-537-2	Sequence 2, Appli	222	36	33.0	35	4	US-09-350-641C-139	Sequence 139, App
150	37	33.9	1403	4	US-09-262-537-6	Sequence 6, Appli	223	36	33.0	35	4	US-09-350-641C-140	Sequence 140, App
151	37	33.9	1658	2	US-08-609-049A-13	Sequence 13, Appl	224	36	33.0	35	4	US-09-350-641C-141	Sequence 141, App
152	37	33.9	1658	3	US-09-170-996-13	Sequence 13, Appl	225	36	33.0	35	4	US-09-350-641C-142	Sequence 142, App
153	37	33.9	1866	4	US-09-355-160D-2	Sequence 2, Appli	226	36	33.0	35	4	US-09-350-641C-143	Sequence 143, App
154	37	33.9	1866	4	US-10-092-219-2	Sequence 2, Appli	227	36	33.0	35	4	US-09-350-641C-144	Sequence 144, App
155	37	33.9	1726	2	US-08-609-049A-30	Sequence 30, Appl	228	36	33.0	35	4	US-09-350-641C-145	Sequence 145, App
156	37	33.9	1726	3	US-09-170-996-30	Sequence 30, Appl	229	36	33.0	35	4	US-09-350-641C-146	Sequence 146, App
157	36.5	33.5	451	4	US-09-489-039A-12702	Sequence 12702, A	230	36	33.0	35	4	US-09-350-641C-147	Sequence 147, App
158	36.5	33.5	724	4	US-09-710-279-3082	Sequence 3082, Ap	231	36	33.0	35	4	US-09-350-641C-148	Sequence 148, App
159	36	33.0	28	3	US-09-082-279B-18	Sequence 18, Appl	232	36	33.0	35	4	US-09-350-641C-149	Sequence 149, App
160	36	33.0	28	3	US-08-474-349A-224	Sequence 224, App	233	36	33.0	35	4	US-09-350-841A-22	Sequence 22, Appl
161	36	33.0	28	3	US-09-315-304B-18	Sequence 18, Appl	234	36	33.0	35	4	US-09-350-841A-138	Sequence 138, App
162	36	33.0	28	4	US-09-834-784-18	Sequence 18, Appl	235	36	33.0	35	4	US-09-350-841A-139	Sequence 139, App
163	36	33.0	28	4	US-09-515-965A-18	Sequence 18, Appl	236	36	33.0	35	4	US-09-350-841A-140	Sequence 140, App
164	36	33.0	28	4	US-09-350-641C-18	Sequence 18, Appl	237	36	33.0	35	4	US-09-350-841A-141	Sequence 141, App
165	36	33.0	28	4	US-09-350-841A-18	Sequence 18, Appl	238	36	33.0	35	4	US-09-350-841A-142	Sequence 142, App
166	36	33.0	35	3	US-08-488-264A-210	Sequence 210, App	239	36	33.0	35	4	US-09-350-841A-143	Sequence 143, App
167	36	33.0	35	3	US-09-082-279B-22	Sequence 22, Appl	240	36	33.0	35	4	US-09-350-841A-144	Sequence 144, App
168	36	33.0	35	3	US-09-082-279B-138	Sequence 138, App	241	36	33.0	35	4	US-09-350-841A-145	Sequence 145, App
169	36	33.0	35	3	US-09-082-279B-139	Sequence 139, App	242	36	33.0	35	4	US-09-350-841A-146	Sequence 146, App
170	36	33.0	35	3	US-09-082-279B-140	Sequence 140, App	243	36	33.0	35	4	US-09-350-841A-147	Sequence 147, App
171	36	33.0	35	3	US-09-082-279B-141	Sequence 141, App	244	36	33.0	35	4	US-09-350-841A-148	Sequence 148, App
172	36	33.0	35	3	US-09-082-279B-142	Sequence 142, App	245	36	33.0	35	4	US-09-350-841A-149	Sequence 149, App
173	36	33.0	35	3	US-09-082-279B-143	Sequence 143, App	246	36	33.0	37	3	US-08-919-597-210	Sequence 210, App

247	36	33.0	37	3	US-08-919-597-211	Sequence 211, App	320	36	33.0	401	4	US-09-270-767-56710	Sequence 56710, A
248	36	33.0	37	3	US-08-919-597-212	Sequence 212, App	321	36	33.0	437	1	US-08-487-037-2	Sequence 2, Appli
249	36	33.0	37	4	US-08-470-896-210	Sequence 210, App	322	36	33.0	437	1	US-08-487-037-3	Sequence 3, Appli
250	36	33.0	37	4	US-08-470-896-211	Sequence 211, App	323	36	33.0	448	1	US-08-295-411-3	Sequence 3, Appli
251	36	33.0	37	4	US-08-470-896-212	Sequence 212, App	324	36	33.0	448	1	US-08-955-471-3	Sequence 3, Appli
252	36	33.0	42	3	US-08-486-099-121	Sequence 121, App	325	36	33.0	448	5	PCT-US92-10068-1	Sequence 1, Appli
253	36	33.0	42	3	US-08-484-223B-121	Sequence 121, App	326	36	33.0	448	5	PCT-US92-10242-3	Sequence 3, Appli
254	36	33.0	42	3	US-08-919-597-121	Sequence 121, App	327	36	33.0	475	4	US-09-902-540-15337	Sequence 15337, A
255	36	33.0	42	3	US-08-475-668A-121	Sequence 121, App	328	36	33.0	475	4	US-09-252-991A-26966	Sequence 26966, A
256	36	33.0	42	3	US-08-485-551A-121	Sequence 121, App	329	36	33.0	488	1	US-08-487-037-1	Sequence 1, Appli
257	36	33.0	42	3	US-08-471-913A-121	Sequence 121, App	330	36	33.0	488	4	US-09-367-777-44	Sequence 44, Appli
258	36	33.0	42	3	US-08-485-264A-121	Sequence 121, App	331	36	33.0	488	4	US-09-367-791A-27	Sequence 27, Appli
259	36	33.0	42	3	US-09-082-279B-17	Sequence 17, Appli	332	36	33.0	496	4	US-09-949-016-9524	Sequence 9524, Ap
260	36	33.0	42	3	US-08-474-349A-121	Sequence 121, App	333	36	33.0	521	4	US-09-252-991A-27321	Sequence 27321, A
261	36	33.0	42	3	US-09-315-304B-17	Sequence 17, Appli	334	36	33.0	527	2	US-08-659-939-4	Sequence 4, Appli
262	36	33.0	42	4	US-08-470-896-121	Sequence 121, App	335	36	33.0	527	2	US-08-850-041-4	Sequence 4, Appli
263	36	33.0	42	4	US-08-485-546A-121	Sequence 121, App	336	36	33.0	527	3	US-08-896-500-4	Sequence 4, Appli
264	36	33.0	42	4	US-08-474-784-17	Sequence 17, Appli	337	36	33.0	527	3	US-08-476-397-4	Sequence 4, Appli
265	36	33.0	42	4	US-09-515-965A-17	Sequence 17, Appli	338	36	33.0	527	3	US-08-973-720-4	Sequence 4, Appli
266	36	33.0	42	4	US-09-350-641C-17	Sequence 17, Appli	339	36	33.0	527	3	US-08-923-558-2	Sequence 2, Appli
267	36	33.0	42	4	US-09-350-841A-17	Sequence 17, Appli	340	36	33.0	527	3	US-09-262-927-4	Sequence 2, Appli
268	36	33.0	42	4	US-08-487-266A-121	Sequence 121, App	341	36	33.0	527	3	US-09-486-553-2	Sequence 2, Appli
269	36	33.0	48	3	US-08-486-099-16	Sequence 16, Appli	342	36	33.0	527	4	US-09-570-383-4	Sequence 4, Appli
270	36	33.0	48	3	US-08-360-107A-16	Sequence 16, Appli	343	36	33.0	527	4	US-09-395-336-4	Sequence 4, Appli
271	36	33.0	48	3	US-08-484-223B-16	Sequence 16, Appli	344	36	33.0	534	4	US-09-252-991A-32086	Sequence 32086, A
272	36	33.0	48	3	US-08-919-597-16	Sequence 16, Appli	345	36	33.0	569	1	US-08-306-231-3	Sequence 3, Appli
273	36	33.0	48	3	US-08-475-668A-16	Sequence 16, Appli	346	36	33.0	574	6	5223254-2	Patent No. 5223254
274	36	33.0	48	3	US-08-485-551A-16	Sequence 16, Appli	347	36	33.0	574	6	5223254-2	Patent No. 5223254
275	36	33.0	48	3	US-08-471-913A-16	Sequence 16, Appli	348	36	33.0	592	4	US-09-919-039-141	Sequence 141, App
276	36	33.0	48	3	US-08-485-264A-16	Sequence 16, Appli	349	36	33.0	594	2	US-08-659-939-2	Sequence 2, Appli
277	36	33.0	48	3	US-08-474-349A-16	Sequence 16, Appli	350	36	33.0	594	2	US-08-850-041-2	Sequence 2, Appli
278	36	33.0	48	4	US-08-255-208A-33	Sequence 33, Appli	351	36	33.0	594	2	US-08-467-963C-6	Sequence 6, Appli
279	36	33.0	48	4	US-08-470-896-16	Sequence 16, Appli	352	36	33.0	594	2	US-08-838-189D-6	Sequence 6, Appli
280	36	33.0	48	4	US-08-485-546A-16	Sequence 16, Appli	353	36	33.0	594	3	US-08-852-344D-6	Sequence 6, Appli
281	36	33.0	48	4	US-08-487-266A-16	Sequence 16, Appli	354	36	33.0	594	3	US-08-896-500-2	Sequence 2, Appli
282	36	33.0	50	4	US-08-255-208A-16	Sequence 16, Appli	355	36	33.0	594	3	US-08-476-397-2	Sequence 2, Appli
283	36	33.0	127	3	US-08-946-026-4	Sequence 4, Appli	356	36	33.0	594	3	US-08-973-720-2	Sequence 2, Appli
284	36	33.0	141	4	US-09-732-210-1657	Sequence 1657, App	357	36	33.0	594	3	US-08-344-6395E-6	Sequence 6, Appli
285	36	33.0	142	4	US-09-732-210-519	Sequence 519, App	358	36	33.0	594	3	US-09-262-927-2	Sequence 2, Appli
286	36	33.0	159	4	US-09-248-796A-19824	Sequence 19824, A	359	36	33.0	594	3	US-08-467-969A-6	Sequence 6, Appli
287	36	33.0	167	4	US-09-270-767-45525	Sequence 45525, A	360	36	33.0	594	3	US-08-467-961A-6	Sequence 6, Appli
288	36	33.0	171	4	US-09-248-796A-19456	Sequence 19456, A	361	36	33.0	594	3	US-08-001-554A-6	Sequence 6, Appli
289	36	33.0	180	4	US-09-902-540-13416	Sequence 13416, A	362	36	33.0	594	4	US-09-570-383-2	Sequence 2, Appli
290	36	33.0	235	3	US-09-029-785-1	Sequence 1, Appli	363	36	33.0	594	4	US-09-395-336-2	Sequence 2, Appli
291	36	33.0	235	4	US-08-382-430-3	Sequence 3, Appli	364	36	33.0	632	4	US-09-949-016-10668	Sequence 10668, A
292	36	33.0	241	1	US-08-330-978-4	Sequence 4, Appli	365	36	33.0	637	4	US-09-949-016-8152	Sequence 8152, Ap
293	36	33.0	241	1	US-08-474-042-4	Sequence 4, Appli	366	36	33.0	643	4	US-09-538-092-844	Sequence 844, App
294	36	33.0	241	1	US-08-484-558-4	Sequence 4, Appli	367	36	33.0	651	4	US-09-248-796A-18743	Sequence 18743, A
295	36	33.0	241	1	US-08-774-592-4	Sequence 4, Appli	368	36	33.0	668	4	US-09-270-767-41490	Sequence 41490, A
296	36	33.0	247	3	US-08-944-483-49	Sequence 49, Appli	369	36	33.0	681	6	5194595-19	Patent No. 5194595
297	36	33.0	252	4	US-09-248-796A-18740	Sequence 18740, A	370	36	33.0	681	6	5194595-19	Patent No. 5194595
298	36	33.0	254	1	US-08-330-978-3	Sequence 3, Appli	371	36	33.0	749	4	US-09-562-737-95	Sequence 95, Appli
299	36	33.0	254	1	US-08-474-042-3	Sequence 3, Appli	372	36	33.0	783	4	US-08-780-562-7	Sequence 7, Appli
300	36	33.0	254	1	US-08-484-558-3	Sequence 3, Appli	373	36	33.0	804	4	US-09-116-676-10	Sequence 10, Appli
301	36	33.0	254	1	US-08-774-592-3	Sequence 3, Appli	374	36	33.0	819	4	US-09-949-016-11044	Sequence 11044, A
302	36	33.0	272	1	US-08-282-951-6	Sequence 6, Appli	375	36	33.0	837	4	US-09-949-016-6515	Sequence 6515, Ap
303	36	33.0	280	4	US-09-601-478-5	Sequence 5, Appli	376	36	33.0	883	4	US-08-982-430-1	Sequence 1, Appli
304	36	33.0	280	4	US-09-601-478-8	Sequence 8, Appli	377	36	33.0	894	2	US-08-599-455B-2	Sequence 2, Appli
305	36	33.0	280	4	US-09-442-013-13	Sequence 13, Appli	378	36	33.0	894	3	US-09-069-781B-2	Sequence 2, Appli
306	36	33.0	280	4	US-09-513-365A-1	Sequence 1, Appli	379	36	33.0	894	3	US-08-618-957A-12	Sequence 12, Appli
307	36	33.0	280	4	US-09-513-365A-6	Sequence 6, Appli	380	36	33.0	894	3	US-09-137-132-2	Sequence 2, Appli
308	36	33.0	306	1	US-08-330-978-1	Sequence 1, Appli	381	36	33.0	894	3	US-08-864-564A-2	Sequence 2, Appli
309	36	33.0	306	1	US-08-474-042-1	Sequence 1, Appli	382	36	33.0	894	4	US-09-094-410-2	Sequence 2, Appli
310	36	33.0	306	1	US-08-484-558-1	Sequence 1, Appli	383	36	33.0	894	4	US-08-708-123D-2	Sequence 2, Appli
311	36	33.0	306	1	US-08-774-592-1	Sequence 1, Appli	384	36	33.0	894	4	US-08-583-153A-2	Sequence 2, Appli
312	36	33.0	316	4	US-09-328-352-8163	Sequence 8163, Ap	385	36	33.0	894	4	US-08-570-142D-2	Sequence 2, Appli
313	36	33.0	320	4	US-09-252-991A-21360	Sequence 21360, A	386	36	33.0	894	4	US-08-638-524B-2	Sequence 2, Appli
314	36	33.0	325	4	US-09-252-991A-32561	Sequence 32561, A	387	36	33.0	894	4	US-10-095-929-12	Sequence 12, Appli
315	36	33.0	326	4	US-09-543-681A-4563	Sequence 4563, Ap	388	36	33.0	895	3	US-08-827-962-19	Sequence 19, Appli
316	36	33.0	349	4	US-09-902-540-13240	Sequence 13240, A	389	36	33.0	895	3	US-08-827-962-21	Sequence 21, Appli
317	36	33.0	359	4	US-09-972-758A-2	Sequence 2, Appli	390	36	33.0	896	2	US-08-640-389A-10	Sequence 10, Appli
318	36	33.0	362	4	US-09-252-991A-24409	Sequence 24409, A	391	36	33.0	896	2	US-08-640-389A-12	Sequence 12, Appli
319	36	33.0	367	4	US-09-252-991A-28477	Sequence 28477, A	392	36	33.0	896	3	US-08-618-957A-10	Sequence 10, Appli

393	36	33.0	896	4	US-09-357-914-33	Sequence 33, Appl	466	35	32.1	132	4	US-09-732-210-1658	Sequence 1658, Ap
394	36	33.0	896	4	US-08-780-562-3	Sequence 3, Appli	467	35	32.1	142	4	US-09-270-767-37079	Sequence 37079, A
395	36	33.0	896	4	US-10-095-929-10	Sequence 10, Appl	468	35	32.1	142	4	US-09-270-767-52296	Sequence 52296, A
396	36	33.0	898	2	US-08-693-697-36	Sequence 36, Appl	469	35	32.1	143	4	US-09-152-060-107	Sequence 107, App
397	36	33.0	898	2	US-08-588-189-3	Sequence 3, Appli	470	35	32.1	145	4	US-09-270-767-39255	Sequence 39255, A
398	36	33.0	906	2	US-08-640-389A-9	Sequence 9, Appli	471	35	32.1	145	4	US-09-270-767-54472	Sequence 54472, A
399	36	33.0	906	3	US-08-618-957A-9	Sequence 9, Appli	472	35	32.1	158	4	US-09-248-796A-18178	Sequence 18178, A
400	36	33.0	906	4	US-09-357-914-32	Sequence 32, Appl	473	35	32.1	159	4	US-09-713-273A-6	Sequence 6, Appli
401	36	33.0	906	4	US-10-095-929-9	Sequence 9, Appli	474	35	32.1	168	4	US-09-270-767-37245	Sequence 37245, A
402	36	33.0	908	2	US-08-693-697-33	Sequence 33, Appl	475	35	32.1	168	4	US-09-270-767-52462	Sequence 52462, A
403	36	33.0	908	2	US-08-588-526-3	Sequence 3, Appli	476	35	32.1	172	4	US-09-252-991A-22192	Sequence 22192, A
404	36	33.0	923	4	US-08-780-562-4	Sequence 4, Appli	477	35	32.1	177	4	US-09-248-796A-20492	Sequence 20492, A
405	36	33.0	937	4	US-09-252-991A-19446	Sequence 19446, A	478	35	32.1	188	3	US-08-858-207A-445	Sequence 445, App
406	36	33.0	958	2	US-08-640-389A-8	Sequence 8, Appli	479	35	32.1	188	4	US-09-583-110-3036	Sequence 3036, Ap
407	36	33.0	958	3	US-08-618-957A-8	Sequence 8, Appli	480	35	32.1	194	4	US-09-107-433-2363	Sequence 2363, Ap
408	36	33.0	958	4	US-10-095-929-8	Sequence 8, Appli	481	35	32.1	214	2	US-08-846-790A-1	Sequence 1, Appli
409	36	33.0	960	1	US-08-355-888A-8	Sequence 8, Appli	482	35	32.1	214	3	US-08-935-333-1	Sequence 1, Appli
410	36	33.0	960	2	US-08-588-190-3	Sequence 3, Appli	483	35	32.1	229	4	US-09-252-991A-22195	Sequence 22195, A
411	36	33.0	960	2	US-08-693-697-8	Sequence 8, Appli	484	35	32.1	236	4	US-09-248-796A-22126	Sequence 22126, A
412	36	33.0	960	2	US-08-640-389A-3	Sequence 3, Appli	485	35	32.1	260	2	US-09-193-877-3	Sequence 3, Appli
413	36	33.0	960	3	US-08-693-696-8	Sequence 8, Appli	486	35	32.1	281	3	US-08-808-148-4	Sequence 4, Appli
414	36	33.0	960	3	US-08-618-957A-3	Sequence 3, Appli	487	35	32.1	289	4	US-09-583-110-3563	Sequence 3563, Ap
415	36	33.0	960	4	US-09-357-914-8	Sequence 8, Appli	488	35	32.1	292	4	US-09-107-433-4868	Sequence 4868, Ap
416	36	33.0	960	4	US-10-095-929-3	Sequence 3, Appli	489	35	32.1	293	4	US-09-248-796A-18514	Sequence 18514, A
417	36	33.0	1162	2	US-08-599-455B-43	Sequence 43, Appl	490	35	32.1	299	4	US-09-270-767-43349	Sequence 43349, A
418	36	33.0	1162	3	US-08-827-963-15	Sequence 15, Appl	491	35	32.1	303	4	US-09-248-796A-16452	Sequence 16452, A
419	36	33.0	1162	3	US-08-827-963-20	Sequence 20, Appl	492	35	32.1	334	3	US-09-267-031-2	Sequence 2, Appli
420	36	33.0	1162	3	US-08-803-346-1	Sequence 1, Appli	493	35	32.1	337	2	US-08-861-464-12	Sequence 12, Appl
421	36	33.0	1162	3	US-09-069-781B-43	Sequence 43, Appl	494	35	32.1	337	2	US-08-396-001-12	Sequence 12, Appl
422	36	33.0	1162	3	US-09-137-132-43	Sequence 43, Appl	495	35	32.1	337	3	US-09-323-433A-12	Sequence 12, Appl
423	36	33.0	1162	3	US-08-864-564A-43	Sequence 43, Appl	496	35	32.1	337	4	US-09-826-752-12	Sequence 12, Appl
424	36	33.0	1162	4	US-09-094-410-43	Sequence 43, Appl	497	35	32.1	358	4	US-09-107-433-3157	Sequence 3157, Ap
425	36	33.0	1162	4	US-08-708-123D-43	Sequence 43, Appl	498	35	32.1	365	1	US-08-447-500-4	Sequence 4, Appli
426	36	33.0	1162	4	US-08-638-524B-43	Sequence 43, Appl	499	35	32.1	365	1	US-08-454-097-4	Sequence 4, Appli
427	36	33.0	1165	2	US-08-640-389A-11	Sequence 11, Appl	500	35	32.1	365	1	US-08-453-866-4	Sequence 4, Appli
428	36	33.0	1165	2	US-08-599-455B-4	Sequence 4, Appli							
429	36	33.0	1165	3	US-09-093-814-1	Sequence 1, Appli							
430	36	33.0	1165	3	US-09-069-781B-4	Sequence 4, Appli							
431	36	33.0	1165	3	US-08-618-957A-11	Sequence 11, Appl							
432	36	33.0	1165	3	US-09-137-132-4	Sequence 4, Appli							
433	36	33.0	1165	3	US-08-864-564A-4	Sequence 4, Appli							
434	36	33.0	1165	4	US-09-094-410-4	Sequence 4, Appli							
435	36	33.0	1165	4	US-08-708-123D-4	Sequence 4, Appli							
436	36	33.0	1165	4	US-08-583-153A-4	Sequence 4, Appli							
437	36	33.0	1165	4	US-08-570-142D-4	Sequence 4, Appli							
438	36	33.0	1165	4	US-08-780-562-2	Sequence 2, Appli							
439	36	33.0	1165	4	US-08-638-524B-4	Sequence 4, Appli							
440	36	33.0	1165	4	US-10-095-929-11	Sequence 11, Appl							
441	36	33.0	1436	4	US-09-578-063-78	Sequence 78, Appl							
442	36	33.0	3200	2	US-08-477-451-8	Sequence 8, Appli							
443	36	33.0	5179	4	US-09-538-092-1258	Sequence 1258, Ap							
444	35.5	32.6	139	4	US-09-902-540-15253	Sequence 15253, A							
445	35.5	32.6	162	3	US-09-199-637A-162	Sequence 162, App							
446	35.5	32.6	175	4	US-09-949-016-10743	Sequence 10743, A							
447	35.5	32.6	179	4	US-09-252-991A-30153	Sequence 30153, A							
448	35.5	32.6	269	4	US-10-000-489-78	Sequence 78, Appl							
449	35.5	32.6	269	4	US-09-949-016-6121	Sequence 6121, Ap							
450	35.5	32.6	276	4	US-09-949-016-7261	Sequence 7261, Ap							
451	35.5	32.6	505	3	US-08-426-509A-17	Sequence 17, Appl							
452	35.5	32.6	505	4	US-08-232-545-17	Sequence 17, Appl							
453	35.5	32.6	505	5	PCT-US95-05008-17	Sequence 17, Appl							
454	35.5	32.6	573	4	US-09-949-016-10415	Sequence 10415, A							
455	35.5	32.6	636	4	US-09-252-991A-24902	Sequence 24902, A							
456	35	32.1	77	4	US-09-152-060-67	Sequence 67, Appl							
457	35	32.1	93	2	US-08-479-078-11	Sequence 11, Appl							
458	35	32.1	93	4	US-09-538-092-139	Sequence 139, App							
459	35	32.1	98	1	US-09-198-452A-1291	Sequence 1291, Ap							
460	35	32.1	99	1	US-08-473-981A-12	Sequence 12, Appl							
461	35	32.1	99	2	US-08-474-087-12	Sequence 12, Appl							
462	35	32.1	100	4	US-09-621-976-4269	Sequence 4269, Ap							
463	35	32.1	118	4	US-09-489-039A-7782	Sequence 7782, Ap							
464	35	32.1	130	4	US-09-732-210-1645	Sequence 1645, Ap							
465	35	32.1	132	4	US-09-732-210-1655	Sequence 1655, Ap							

ALIGNMENTS

RESULT 1

US-09-360-376-55

; Sequence 55, Application US/09360376

; Patent No. 6495739

; GENERAL INFORMATION:

; APPLICANT: Lasener, Michael

; APPLICANT: Ruesinsky, Diane

; TITLE OF INVENTION: PLANT PHOSPHATIDIC ACID PHOSPHATASES

; FILE REFERENCE: 17026/01/US

; CURRENT APPLICATION NUMBER: US/09/360,376

; CURRENT FILING DATE: 1999-07-23

; PRIOR APPLICATION NUMBER: US 09/122,315

; PRIOR FILING DATE: 1998-07-24

; NUMBER OF SEQ ID NOS: 55

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 55

; LENGTH: 233

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-360-376-55

Query Match 100.0%; Score 109; DB 4; Length 233;

Best Local Similarity 100.0%; Pred. No. 1.5e-09;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGYIONVRCRGDSKVQEAR 20

DB 173 EGYIONVRCRGDSKVQEAR 192

RESULT 2

US-09-360-376-54

; Sequence 54, Application US/09360376


```
; Patent No. 6495739
; GENERAL INFORMATION:
; APPLICANT: Lassner, Michael
; APPLICANT: Ruezinsky, Diane
; TITLE OF INVENTION: PLANT PHOSPHATIDIC ACID PHOSPHATASES
; FILE REFERENCE: 17026/01/US
; CURRENT APPLICATION NUMBER: US/09/360,376
; CURRENT FILING DATE: 1999-07-23
; PRIOR APPLICATION NUMBER: US 09/122,315
; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 54
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-360-376-54

Query Match          60.6%; Score 66; DB 4; Length 282;
Best Local Similarity 50.0%; Pred. No. 0.012;
Matches 10; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY      1 EGYIQRRCRGDDSKVQEAR 20
       :|||:|:|:|:|:|:|:|:|
Db      145 DGYIENFVCGNEQKVEGR 164

RESULT 3
US-08-992-035A-3
; Sequence 3, Application US/08992035A
; Patent No. 6242179
; GENERAL INFORMATION:
; APPLICANT: Shah, Purvi
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: HUMAN PHOSPHATASES
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/992,035A
; FILING DATE: December 17, 1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0433 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 283 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 1487873
US-08-992-035A-3
```

```
Query Match          59.6%; Score 65; DB 3; Length 283;
Best Local Similarity 50.0%; Pred. No. 0.017;
Matches 10; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY      1 EGYIQRRCRGDDSKVQEAR 20
       :|||:|:|:|:|:|:|:|:|
Db      145 DGYIEDYICQNEEKVKEGR 164

RESULT 4
US-09-360-376-53
; Sequence 53, Application US/09360376
; Patent No. 6495739
; GENERAL INFORMATION:
; APPLICANT: Lassner, Michael
; APPLICANT: Ruezinsky, Diane
; TITLE OF INVENTION: PLANT PHOSPHATIDIC ACID PHOSPHATASES
; FILE REFERENCE: 17026/01/US
; CURRENT APPLICATION NUMBER: US/09/360,376
; CURRENT FILING DATE: 1999-07-23
; PRIOR APPLICATION NUMBER: US 09/122,315
; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 53
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-360-376-53

Query Match          59.6%; Score 65; DB 4; Length 283;
Best Local Similarity 50.0%; Pred. No. 0.017;
Matches 10; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY      1 EGYIQRRCRGDDSKVQEAR 20
       :|||:|:|:|:|:|:|:|:|
Db      145 DGYIEDYICQNEEKVKEGR 164

RESULT 5
US-08-992-035A-1
; Sequence 1, Application US/08992035A
; Patent No. 6242179
; GENERAL INFORMATION:
; APPLICANT: Shah, Purvi
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: HUMAN PHOSPHATASES
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/992,035A
; FILING DATE: December 17, 1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0433 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 283 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 1487873
US-08-992-035A-3
```

TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 285 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BLADNOT06
CLONE: 1719418
US-08-992-035A-1

Query Match 54.1%; Score 59; DB 3; Length 285;
Best Local Similarity 50.0%; Pred. No. 0.16;
Matches 10; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 EGYQNYRCRGDDSKVQEAR 20
146 DGVEIYYICRGNAERVKEGR 165

RESULT 6
US-08-311-731A-4
Sequence 4, Application US/08311731A
Patent No. 6583266

GENERAL INFORMATION:
APPLICANT: SMITH, DOUGLAS

APPLICANT: MAO, JEN-I

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
RELATION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 411

CORRESPONDENCE ADDRESS:

ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.

STREET: 600 ATLANTIC AVENUE

CITY: BOSTON

STATE: MASSACHUSETTS

COUNTRY: USA

ZIP: 02210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/311,731A

FILING DATE:

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: GATES, EDWARD R.

REGISTRATION NUMBER: 31,616

REFERENCE/DOCKET NUMBER: C0044/7125

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/720-3500

TELEFAX: 617/720-2441

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 705 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: YES

ORIGINAL SOURCE:

ORGANISM: MYCOBACTERIUM TUBERCULOSIS

US-08-311-731A-4

Query Match 42.7%; Score 46.5; DB 4; Length 705;
Best Local Similarity 50.0%; Pred. No. 40;
Matches 8; Conservative 5; Mismatches 2; Indels 1; Gaps 1;

QY 1 EGYQNYRCRGDDSKV 16

DB 378 DGMV-TYRCRADDTGV 392
:|:: |||| |::|

RESULT 7

US-09-949-016-8237

Sequence 8237, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 8237

LENGTH: 416

TYPE: PRT

ORGANISM: Human

US-09-949-016-8237

Query Match 40.4%; Score 44; DB 4; Length 416;
Best Local Similarity 53.8%; Pred. No. 56;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 YIQNYRCRGDDSK 15

:|:: |||| |::|

DB 397 YLYNWRCPGTDSE 409

RESULT 8

US-09-786-240-1

Sequence 1, Application US/09786240

Patent No. 6558935

GENERAL INFORMATION:

APPLICANT: INCYTE PHARMACEUTICALS, INC.

APPLICANT: TANG, Y. Tom

APPLICANT: CORLEY, Neil C.

APPLICANT: GUEGLER, Karl J.

APPLICANT: BAUGHN, Mariah R.

APPLICANT: LAL, Preeti

APPLICANT: YUE, Henry

APPLICANT: HILLMAN, Jennifer L.

APPLICANT: AZIMZAI, Valda

TITLE OF INVENTION: HUMAN TRANSFERASE PROTEINS

FILE REFERENCE: PF-0592 PCT

CURRENT APPLICATION NUMBER: US/09/786,240

CURRENT FILING DATE: 2002-03-12

PRIOR APPLICATION NUMBER: 09/150,657; unassigned; 09/186,779; unassigned; 60/133,642

PRIOR FILING DATE: 1998-09-10; 1998-09-10; 1998-11-04; 1998-11-04; 1999-05-11

NUMBER OF SEQ ID NOS: 33

SOFTWARE: PERL Program

SEQ ID NO 1

LENGTH: 498

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc.feature

OTHER INFORMATION: Incyte ID No. 6558935 1632930CDI

US-09-786-240-1

Query Match 40.4%; Score 44; DB 4; Length 498;
Best Local Similarity 53.8%; Pred. No. 68;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;


```
/ ; STATE: Colorado
/ ; COUNTRY: USA
/ ; ZIP: 80525
/ ; COMPUTER READABLE FORM:
/ ; MEDIUM TYPE: Floppy disk
/ ; COMPUTER: IBM PC compatible
/ ; OPERATING SYSTEM: PC-DOS/MS-DOS
/ ; SOFTWARE: WordPerfect for Windows, Version 7.0
/ ; CURRENT APPLICATION DATA:
/ ; APPLICATION NUMBER: US/09/015,734
/ ; FILING DATE: 29-JAN-1998
/ ; CLASSIFICATION:
/ ; ATTORNEY/AGENT INFORMATION:
/ ; NAME: Verser, Carol Talkington
/ ; REGISTRATION NUMBER: 37,459
/ ; REFERENCE/DOCKET NUMBER: DI-4
/ ; TELECOMMUNICATION INFORMATION:
/ ; TELEPHONE: 970/493-7272
/ ; TELEFAX: 970/484-9505
/ ; INFORMATION FOR SEQ ID NO: 12:
/ ; SEQUENCE CHARACTERISTICS:
/ ; LENGTH: 201 amino acids
/ ; TYPE: amino acid
/ ; TOPOLOGY: linear
/ ; MOLECULE TYPE: Protein
/ ; US-09-015-734-12

Query Match 37.6%; Score 41; DB 3; Length 201;
Best Local Similarity 53.8%; Pred. No. 78;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 7 YRCRGDDSKVQEA 19
Db 91 YRCRNDLNLSEA 103

RESULT 14
US-09-515-311-12
; Sequence 12, Application US/09515311
; Patent No. 6582701
; GENERAL INFORMATION:
; APPLICANT: Weber, Eric R.
; APPLICANT: McCall, Catherine A.
; TITLE OF INVENTION: NOVEL EQUINE FC EPSILON RECEPTOR ALPHA
; TITLE OF INVENTION: CHAIN NUCLEIC ACID MOLECULES, PROTEINS AND USES THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/515,311
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/015,734
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: DI-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 12:
```

```
/ ; SEQUENCE CHARACTERISTICS:
/ ; LENGTH: 201 amino acids
/ ; TYPE: amino acid
/ ; TOPOLOGY: linear
/ ; MOLECULE TYPE: Protein
/ ; US-09-515-311-12

Query Match 37.6%; Score 41; DB 4; Length 201;
Best Local Similarity 53.8%; Pred. No. 78;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 7 YRCRGDDSKVQEA 19
Db 91 YRCRNDLNLSEA 103

RESULT 15
US-09-015-734-7
; Sequence 7, Application US/09015734
; Patent No. 6057127
; GENERAL INFORMATION:
; APPLICANT: Weber, Eric R.
; APPLICANT: McCall, Catherine A.
; TITLE OF INVENTION: NOVEL EQUINE FC EPSILON RECEPTOR ALPHA
; TITLE OF INVENTION: CHAIN NUCLEIC ACID MOLECULES, PROTEINS AND USES THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/015,734
; FILING DATE: 29-JAN-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: DI-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 236 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; US-09-015-734-7

Query Match 37.6%; Score 41; DB 3; Length 236;
Best Local Similarity 53.8%; Pred. No. 92;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 7 YRCRGDDSKVQEA 19
Db 72 YRCRNDLNLSEA 84

RESULT 16
US-09-515-311-7
; Sequence 7, Application US/09515311
; Patent No. 6582701
; GENERAL INFORMATION:
; APPLICANT: Weber, Eric R.
; APPLICANT: McCall, Catherine A.
```

```

; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: DI-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
US-09-015-734-2

Query Match 37.6%; Score 41; DB 3; Length 255;
Best Local Similarity 53.8%; Pred.No.1e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 7 YRCRGDDSKVQEA 19
|||||:|:|
Db 91 YRCRNNDNLSEA 103

RESULT 18
US-09-515-311-2
; Sequence 2, Application US/09515311
; Patent No. 6582701
; GENERAL INFORMATION:
; APPLICANT: Weber, Eric R.
; APPLICANT: McCall, Catherine A.
; TITLE OF INVENTION: NOVEL EQUINE FC EPSILON RECEPTOR ALPHA
; TITLE OF INVENTION: CHAIN NUCLEIC ACID MOLECULES, PROTEINS AND USES THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/515,311
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/015,734
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: DI-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
US-09-515-311-2

Query Match 37.6%; Score 41; DB 4; Length 255;
Best Local Similarity 53.8%; Pred.No.1e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 7 YRCRGDDSKVQEA 19
|||||:|:|

```

```
Db 91 YRCRNDLNLSEA 103

RESULT 19
US-09-489-039A-13748
; Sequence 13748, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13748
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13748

Query Match 37.6%; Score 41; DB 4; Length 272;
Best Local Similarity 50.0%; Pred. NO. 1.1e+02;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 6 NYRCRGDDSKVQEA 19
Db 210 NVRFKGGDDAKAEKA 223

RESULT 20
US-09-506-286B-92
; Sequence 92, Application US/09506286B
; Patent No. 6482414
; GENERAL INFORMATION:
; APPLICANT: Dowling, Patricia W.
; APPLICANT: Youngner, Julius S.
; APPLICANT: The University of Pittsburgh, of the Commonwealth
; TITLE OF INVENTION: COLD-ADAPTED EQUINE INFLUENZA VIRUSES
; FILE REFERENCE: EQ-1-C2
; CURRENT APPLICATION NUMBER: US/09/506,286B
; CURRENT FILING DATE: 2000-02-16
; PRIOR APPLICATION NUMBER: 09/133,921
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: PCT/US99/18583
; PRIOR FILING DATE: 1999-08-12
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 92
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Equine influenza virus H3N8
US-09-506-286B-92

Query Match 37.6%; Score 41; DB 4; Length 396;
Best Local Similarity 38.9%; Pred. NO. 1.6e+02;
Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 3 YIQNYRCRGDDSKVQEAR 20
Db 194 YRYTYRCQRGDTQIQTRR 211

RESULT 21
US-09-506-286B-95
; Sequence 95, Application US/09506286B
; Patent No. 6482414
; GENERAL INFORMATION:
; APPLICANT: Dowling, Patricia W.
; APPLICANT: Youngner, Julius S.
; APPLICANT: The University of Pittsburgh, of the Commonwealth
; TITLE OF INVENTION: COLD-ADAPTED EQUINE INFLUENZA VIRUSES
; FILE REFERENCE: EQ-1-C2-1
; CURRENT APPLICATION NUMBER: US/09/506,286B
; CURRENT FILING DATE: 2000-02-16
; PRIOR APPLICATION NUMBER: 09/133,921
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: PCT/US99/18583
; PRIOR FILING DATE: 1999-08-12
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 95
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Equine influenza virus H3N8
US-09-506-286B-95

Query Match 37.6%; Score 41; DB 4; Length 396;
Best Local Similarity 38.9%; Pred. NO. 1.6e+02;
Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 3 YIQNYRCRGDDSKVQEAR 20
Db 194 YRYTYRCQRGDTQIQTRR 211

RESULT 22
US-10-065-133A-92
; Sequence 92, Application US/10065133A
; Patent No. 6685946
; GENERAL INFORMATION:
; APPLICANT: Dowling, Patricia W.
; APPLICANT: Youngner, Julius S.
; APPLICANT: The University of Pittsburgh, of the Commonwealth
; TITLE OF INVENTION: COLD-ADAPTED EQUINE INFLUENZA VIRUSES
; FILE REFERENCE: EQ-1-C2-1
; CURRENT APPLICATION NUMBER: US/10/065,133A
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: PCT/US99/18583
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 09/133,921
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 92
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Equine influenza virus H3N8
US-10-065-133A-92

Query Match 37.6%; Score 41; DB 4; Length 396;
Best Local Similarity 38.9%; Pred. NO. 1.6e+02;
Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 3 YIQNYRCRGDDSKVQEAR 20
Db 194 YRYTYRCQRGDTQIQTRR 211

RESULT 23
US-10-065-133A-95
; Sequence 95, Application US/10065133A
; Patent No. 6685946
; GENERAL INFORMATION:
; APPLICANT: Dowling, Patricia W.
; APPLICANT: Youngner, Julius S.
; APPLICANT: The University of Pittsburgh, of the Commonwealth
; TITLE OF INVENTION: COLD-ADAPTED EQUINE INFLUENZA VIRUSES
; FILE REFERENCE: EQ-1-C2-1
; CURRENT APPLICATION NUMBER: US/10/065,133A
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: PCT/US99/18583
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 09/133,921
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 95
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Equine influenza virus H3N8
US-10-065-133A-95

Query Match 37.6%; Score 41; DB 4; Length 396;
Best Local Similarity 38.9%; Pred. NO. 1.6e+02;
Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 3 YIQNYRCRGDDSKVQEAR 20
Db 194 YRYTYRCQRGDTQIQTRR 211
```

```

; NAME/KEY: misc feature
; LOCATION: (707)..(707)
; OTHER INFORMATION: The 'Xaa' at location 707 stands for Arg, or Lys.
US-10-065-133A-107

Query Match          37.6%; Score 41; DB 4; Length 757;
Best Local Similarity 38.9%; Pred. NO. 3.2e+02;
Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY      3 YIQNYRCRGDDSKVQEAR 20
      | ||| : ||| :
Db      555 YRYTRCQRGDTQIQTRR 572
      | ||| : ||| :

RESULT 26
US-09-902-540-13372
; Sequence 13372, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 13372
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-13372

Query Match          37.2%; Score 40.5; DB 4; Length 145;
Best Local Similarity 52.6%; Pred. NO. 66;
Matches 10; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY      2 GYIQNYRCRGDDSKVQEAR 20
      | ||| : ||| :
Db      45 GSIQTWR-NSDEEKVQEMR 62
      | ||| : ||| :

RESULT 27
US-09-252-991A-20336
; Sequence 20336, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20336
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20336

Query Match          37.2%; Score 40.5; DB 4; Length 314;
Best Local Similarity 42.9%; Pred. NO. 1.5e+02;
Matches 9; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY      1 EGYIQNYRC-RGDDSKVQEAR 20
      | ||| : ||| :

```

```
Db 172 DGYVTNLNVRHGDYARVGEAK 192

RESULT 28
US-09-198-452A-1290
; Sequence 1290, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Grifais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1290
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-1290

Query Match 36.7%; Score 40; DB 4; Length 94;
Best Local Similarity 63.6%; Pred. No. 50;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 GYQNYRCRGD 12
Db 9 GYVNRRCRGD 19

RESULT 29
US-09-134-001C-5278
; Sequence 5278, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5278
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5278

Query Match 36.7%; Score 40; DB 3; Length 256;
Best Local Similarity 47.1%; Pred. No. 1.4e+02;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 GYQNYRCRGD 18
Db 87 GYQTRKRKDRGRLE 103

RESULT 30
US-09-506-286B-86
; Sequence 86, Application US/09506286B
; Patent No. 6482414
; GENERAL INFORMATION:
; APPLICANT: Dowling, Patricia W.
; APPLICANT: Youngner, Julius S.
; TITLE OF INVENTION: The University of Pittsburgh, of the Commonwealth
; TITLE OF INVENTION: COLD-ADAPTED EQUINE INFLUENZA VIRUSES
; FILE REFERENCE: EQ-1-C2
; CURRENT APPLICATION NUMBER: US/09/506,286B
```

```
; CURRENT FILING DATE: 2000-02-16
; PRIOR APPLICATION NUMBER: 09/133,921
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: PCT/US99/18583
; PRIOR FILING DATE: 1999-08-12
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 86
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Equine influenza virus H3N8
US-09-506-286B-86

Query Match 36.7%; Score 40; DB 4; Length 396;
Best Local Similarity 38.9%; Pred. No. 2.3e+02;
Matches 7; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 3 YIQNYRCRGD 20
Db 194 YRYTYRCHRGDTQTRR 211

RESULT 31
US-09-506-286B-89
; Sequence 89, Application US/09506286B
; Patent No. 6482414
; GENERAL INFORMATION:
; APPLICANT: Dowling, Patricia W.
; APPLICANT: Youngner, Julius S.
; APPLICANT: The University of Pittsburgh, of the Commonwealth
; TITLE OF INVENTION: COLD-ADAPTED EQUINE INFLUENZA VIRUSES
; FILE REFERENCE: EQ-1-C2
; CURRENT APPLICATION NUMBER: US/09/506,286B
; CURRENT FILING DATE: 2000-02-16
; PRIOR APPLICATION NUMBER: 09/133,921
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: PCT/US99/18583
; PRIOR FILING DATE: 1999-08-12
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 89
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Equine influenza virus H3N8
US-09-506-286B-89

Query Match 36.7%; Score 40; DB 4; Length 396;
Best Local Similarity 38.9%; Pred. No. 2.3e+02;
Matches 7; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 3 YIQNYRCRGD 20
Db 194 YRYTYRCHRGDTQTRR 211

RESULT 32
US-10-065-133A-86
; Sequence 86, Application US/10065133A
; Patent No. 6685946
; GENERAL INFORMATION:
; APPLICANT: Dowling, Patricia W.
; APPLICANT: Youngner, Julius S.
; TITLE OF INVENTION: COLD-ADAPTED EQUINE INFLUENZA VIRUSES
; FILE REFERENCE: EQ-1-C2-1
; CURRENT APPLICATION NUMBER: US/10/065,133A
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: PCT/US99/18583
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 09/133,921
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 86
```


; LENGTH: 396
; TYPE: PRT
; ORGANISM: Equine influenza virus H3N8
US-10-065-133A-86

Query Match 36.7%; Score 40; DB 4; Length 396;
Best Local Similarity 38.9%; Pred. No. 2.3e+02;
Matches 7; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 3 YIQNYRCRGDDSKVQEAR 20
| ||| |:::|
Db 194 YRYTYRCHRGDTQIQTR 211

RESULT 33
US-10-065-133A-89
; Sequence 89, Application US/10065133A
; Patent No. 6685946
; GENERAL INFORMATION:

; APPLICANT: Dowling, Patricia W.
; APPLICANT: Youngner, Julius S.
; TITLE OF INVENTION: COLD-ADAPTED EQUINE INFLUENZA VIRUSES
; FILE REFERENCE: EQ-1-C2-1
; CURRENT APPLICATION NUMBER: US/10/065,133A
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: PCT/US99/18583
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 09/133,921
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 89
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Equine influenza virus H3N8
US-10-065-133A-89

Query Match 36.7%; Score 40; DB 4; Length 396;
Best Local Similarity 38.9%; Pred. No. 2.3e+02;
Matches 7; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 3 YIQNYRCRGDDSKVQEAR 20
| ||| |:::|
Db 194 YRYTYRCHRGDTQIQTR 211

RESULT 34
US-09-489-039A-12030
; Sequence 12030, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12030
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12030

Query Match 36.7%; Score 40; DB 4; Length 424;
Best Local Similarity 85.7%; Pred. No. 2.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 YRCRGDD 13
: |||||
Db 63 FRCRGDD 69

RESULT 35
US-09-270-767-41177
; Sequence 41177, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41177
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-41177

Query Match 36.7%; Score 40; DB 4; Length 501;
Best Local Similarity 60.0%; Pred. No. 3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 7 YRCRGDDSKV 16
|||:|::|
Db 485 YRCKGDSSLI 494

RESULT 36
US-09-270-767-56393
; Sequence 56393, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:

; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 56393
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-56393

Query Match 36.7%; Score 40; DB 4; Length 501;
Best Local Similarity 60.0%; Pred. No. 3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 7 YRCRGDDSKV 16
|||:|::|
Db 485 YRCKGDSSLI 494

RESULT 37
US-09-252-991A-19287
; Sequence 19287, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19287
; LENGTH: 710
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19287

Query Match 36.7%; Score 40; DB 4; Length 710;
Best Local Similarity 38.9%; Pred. No. 4.3e+02;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 EGYQNYRCRGDDSKVQE 18
DB 523 QGIAQGHQRGTDRRLQ 540

RESULT 38
US-08-121-713D-58
; Sequence 58, Application US/08121713D
; Patent No. 5639856
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/121,713D
; FILING DATE: 13-SEP-1993
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELEPHONE: (415)343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 730 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-121-713D-58

Query Match 36.7%; Score 40; DB 1; Length 730;
Best Local Similarity 50.0%; Pred. No. 4.4e+02;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 GYQNYRCRGDD 13
DB 648 GFLFSRRCRGED 659

RESULT 39
US-08-835-268-58

; Sequence 58, Application US/08835268
; Patent No. 5807826
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/835,268
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/121,713
; FILING DATE: 13-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)343-4341
; TELEFAX: (415) 343-4342
; TELEX:

; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 730 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-835-268-58

Query Match 36.7%; Score 40; DB 1; Length 730;
Best Local Similarity 50.0%; Pred. No. 4.4e+02;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 GYQNYRCRGDD 13
DB 648 GFLFSRRCRGED 659

RESULT 40
US-09-060-692-58
; Sequence 58, Application US/09060692
; Patent No. 5935865
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104

```
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,692
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/121,713
; FILING DATE: 13-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 730 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-060-692-58
;
; Query Match 36.7%; Score 40; DB 2; Length 730;
; Best Local Similarity 50.0%; Pred. NO. 4.4e+02;
; Mismatches 2; Indels 0; Gaps 0;
;
;
; QY 2 GYIQNYRCRGDD 13
; Db 648 GFLFSRRCRGDD 659
;
; RESULT 41
; US-08-833-391-58
; Sequence 58, Application US/08833391
; Patent No. 6013781
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,391
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/121,713
; FILING DATE: 13-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
```

```
;
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 730 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-833-391-58
;
; Query Match 36.7%; Score 40; DB 3; Length 730;
; Best Local Similarity 50.0%; Pred. NO. 4.4e+02;
; Mismatches 2; Indels 0; Gaps 0;
;
;
; QY 2 GYIQNYRCRGDD 13
; Db 648 GFLFSRRCRGDD 659
;
; RESULT 42
; US-09-060-610-58
; Sequence 58, Application US/09060610
; Patent No. 6344544
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,610
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/835,268
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 730 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-060-610-58
;
; Query Match 36.7%; Score 40; DB 3; Length 730;
; Best Local Similarity 50.0%; Pred. NO. 4.4e+02;
; Mismatches 2; Indels 0; Gaps 0;
;
;
; QY 2 GYIQNYRCRGDD 13
; Db 648 GFLFSRRCRGDD 659
;
; RESULT 43
; US-09-060-610-58
; Sequence 58, Application US/09060610
; Patent No. 6344544
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,610
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/835,268
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 730 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-060-610-58
;
; Query Match 36.7%; Score 40; DB 3; Length 730;
; Best Local Similarity 50.0%; Pred. NO. 4.4e+02;
; Mismatches 2; Indels 0; Gaps 0;
;
;
; QY 2 GYIQNYRCRGDD 13
; Db 648 GFLFSRRCRGDD 659
;
; RESULT 44
; US-09-060-610-58
; Sequence 58, Application US/09060610
; Patent No. 6344544
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,610
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/835,268
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 730 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-060-610-58
;
; Query Match 36.7%; Score 40; DB 3; Length 730;
; Best Local Similarity 50.0%; Pred. NO. 4.4e+02;
; Mismatches 2; Indels 0; Gaps 0;
;
;
; QY 2 GYIQNYRCRGDD 13
; Db 648 GFLFSRRCRGDD 659
;
; RESULT 45
; US-09-060-610-58
; Sequence 58, Application US/09060610
; Patent No. 6344544
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,610
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/835,268
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
```

Db 648 GFLFSRRCRGED 659

RESULT 43
PCT-US94-10151A-58
; Sequence 58, Application PC/TUS9410151A
; GENERAL INFORMATION:
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR HOBBACH TEST ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 13-SEP-1994
; APPLICATION NUMBER: PCT/US94/10151A
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: FP-58750-PC/RAO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299 FHT UR
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 730 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US94-10151A-58

Query Match 36.7%; Score 40; DB 5; Length 730;
Best Local Similarity 50.0%; Pred. No. 4.4e+02;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
QY 2 GYIQNYRCRGDD 13
|::: |||||
Db 648 GFLFSRRCRGED 659

RESULT 44
US-09-506-286B-104
; Sequence 104, Application US/09506286B
; Patent No. 6482414
; GENERAL INFORMATION:
; APPLICANT: Dowling, Patricia W.
; APPLICANT: Youngner, Julius S.
; TITLE OF INVENTION: of Pittsburgh, of the Commonwealth
; FILE REFERENCE: EQ-1-C2
; CURRENT APPLICATION NUMBER: US/09/506,286B
; CURRENT FILING DATE: 2000-02-16
; PRIOR APPLICATION NUMBER: 09/133,921
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: PCT/US99/18583
; PRIOR FILING DATE: 1999-08-12
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 104
; LENGTH: 757
; TYPE: PRT
; ORGANISM: Equine influenza virus H3N8

; FEATURE:
; OTHER INFORMATION: At amino acid location 489, Xaa = unknown
US-09-506-286B-104

Query Match 36.7%; Score 40; DB 4; Length 757;
Best Local Similarity 38.9%; Pred. No. 4.6e+02;
Matches 7; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
QY 3 YIQNYRCRGDDSKVQEAR 20
|::: |||||
Db 555 YRYTYRCHRGDTQIQTRR 572

RESULT 45
US-10-065-133A-104
; Sequence 104, Application US/10065133A
; Patent No. 6685946
; GENERAL INFORMATION:
; APPLICANT: Dowling, Patricia W.
; APPLICANT: Youngner, Julius S.
; TITLE OF INVENTION: COLD-ADAPTED EQUINE INFLUENZA VIRUSES
; FILE REFERENCE: EQ-1-C2-1
; CURRENT APPLICATION NUMBER: US/10/065,133A
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: PCT/US99/18583
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 09/133,921
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 104
; LENGTH: 757
; TYPE: PRT
; ORGANISM: Equine influenza virus H3N8
; NAME/KEY: misc feature
; LOCATION: (489)..(489)
; OTHER INFORMATION: The 'Xaa' at location 489 stands for Thr or Ser
US-10-065-133A-104

Query Match 36.7%; Score 40; DB 4; Length 757;
Best Local Similarity 38.9%; Pred. No. 4.6e+02;
Matches 7; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
QY 3 YIQNYRCRGDDSKVQEAR 20
|::: |||||
Db 555 YRYTYRCHRGDTQIQTRR 572

RESULT 46
US-09-252-991A-30838
; Sequence 30838, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30838
; LENGTH: 980
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30838
Query Match 36.7%; Score 40; DB 4; Length 980;
Best Local Similarity 43.8%; Pred. No. 6.1e+02;

Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 5 QNYRCRGGDDSKVQEAR 20
| | | | | : | |
Db 901 QRYCGRQGENSVHALR 916

RESULT 47

US-09-621-976-4220
; Sequence 4220, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 4220
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -26..-1
US-09-621-976-4220

Query Match 36.2%; Score 39.5; DB 4; Length 99;
Best Local Similarity 69.2%; Pred. No. 63;
Matches 9; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 8 RCRGGDDSKVQEAR 20
| | | | | : | |
Db 41 RCGGDD-KVKAR 52

RESULT 48

US-09-188-930-174
; Sequence 174, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 174
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Human
US-09-188-930-174

Query Match 36.2%; Score 39.5; DB 3; Length 137;
Best Local Similarity 69.2%; Pred. No. 89;
Matches 9; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 8 RCRGGDDSKVQEAR 20
| | | | | : | |
Db 46 RCGGDD-KVKAR 57

RESULT 49

US-09-312-283C-174

; Sequence 174, Application US/09312283C
; Patent No. 6573095
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated from Skin Cells
; FILE REFERENCE: 11000.1011c2
; CURRENT APPLICATION NUMBER: US/09/312,283C
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 174
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Mouse
US-09-312-283C-174

Query Match 36.2%; Score 39.5; DB 4; Length 137;
Best Local Similarity 69.2%; Pred. No. 89;
Matches 9; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 8 RCRGGDDSKVQEAR 20
| | | | | : | |
Db 46 RCGGDD-KVKAR 57

RESULT 50

US-09-673-395A-354
; Sequence 354, Application US/09673395A
; Patent No. 6620923
; GENERAL INFORMATION:
; APPLICANT: SPECHT, THOMAS
; APPLICANT: HINZMANN, BERND
; APPLICANT: SCHMITT, ARMIN
; APPLICANT: PILARSKY, CHRISTIAN
; APPLICANT: DAHL, EDGAR
; APPLICANT: ROSENTHAL, ANDRE
; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM UTERUS TUMOR TISSUE
; FILE REFERENCE: ALBRE-12
; CURRENT APPLICATION NUMBER: US/09/673,395A
; CURRENT FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 637
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 354
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-673-395A-354

Query Match 36.2%; Score 39.5; DB 4; Length 225;
Best Local Similarity 69.2%; Pred. No. 1.5e+02;
Matches 9; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 8 RCRGGDDSKVQEAR 20
| | | | | : | |
Db 120 RCGGDD-KVKAR 131

Search completed: September 7, 2005, 19:58:02
Job time : 31.1163 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2005, 19:43:14 ; Search time 70.2326 Seconds
(without alignments)
145.824 Million cell updates/sec

Title: US-10-812-238B-2

Perfect score: 109

Sequence: 1 EGYIQNYRCRGDDSKVQEAR 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database : Uniprot 03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	109	100.0	311	1 LPPI_HUMAN	O14495 h lipid pho
2	105	96.3	312	1 LPPI_MOUSE	Q99JY8 mus musculus
3	105	96.3	312	1 LPPI_RAT	P97544 rattus norv
4	105	96.3	312	2 Q6IMX4	Q6IMX4 rattus norv
5	74	67.9	307	2 Q68F34	Q68F34 xenopus lae
6	71	65.1	307	2 Q6AX87	Q6AX87 xenopus lae
7	70	64.2	167	2 Q95L84	Q95L84 oryctolagus
8	66	60.6	282	1 LPPI_RAT	O08564 rattus norv
9	66	60.6	282	2 Q6P766	Q6P766 rattus norv
10	65	59.6	283	1 LPPI_MOUSE	Q61469 mus musculus
11	62	56.9	285	1 LPPI_CAVPO	O88956 cavia porce
12	59	54.1	284	1 LPPI_HUMAN	O14494 homo sapien
13	53	48.6	284	2 Q6GP13	Q6GP13 xenopus lae
14	52	47.7	504	1 Y787_METJA	Q58197 methanococc
15	51	46.8	252	2 Q95ZL6	Q95ZL6 caenorhabdi
16	51	46.8	252	2 Q95ZL7	Q95ZL7 caenorhabdi
17	51	46.8	2736	2 Q7RTC3	Q7RTC3 plasmodium
18	50	45.9	276	1 LPPI_MOUSE	Q9DAX2 mus musculus
19	50	45.9	276	1 LPPI_RAT	Q8K593 rattus norv
20	48	44.0	130	2 Q83ER2	Q83ER2 coxiella bu
21	48	44.0	217	2 Q7R7R7	Q7R7R7 plasmodium
22	48	44.0	269	2 Q7R8M3	Q7R8M3 plasmodium
23	48	44.0	283	2 Q6R997	Q6R997 xenopus tro
24	48	44.0	283	2 Q6PD6	Q6PD6 anophelies g
25	47	43.1	343	2 Q6FYZ2	Q6FYZ2 bartonella
26	46.5	42.7	705	2 Q50468	Q50468 mycobacteri
27	46.5	42.7	705	2 P96283	P96283 mycobacteri
28	46.5	42.7	705	2 Q7TXK7	Q7TXK7 mycobacteri
29	46	42.2	130	2 Q6CZY4	Q6CZY4 erwinia car
30	46	42.2	340	2 Q7YU44	Q7YU44 drosophila
31	46	42.2	340	2 Q9VNT9	Q9VNT9 drosophila

32	46	42.2	776	2	Q73RY0	Q73RY0 mycobacteri
33	46	42.2	1325	2	Q95Z20	Q95Z20 caenorhabdi
34	45.5	41.7	179	2	Q20683	Q20683 caenorhabdi
35	45.5	41.7	1086	2	Q6XNK7	Q6XNK7 emericella
36	45	41.3	130	2	Q7MYG5	Q7MYG5 photorhabdu
37	45	41.3	300	2	Q7RA96	Q7RA96 plasmodium
38	45	41.3	304	2	Q7RCS6	Q7RCS6 plasmodium
39	45	41.3	314	2	Q7RS76	Q7RS76 plasmodium
40	45	41.3	344	2	Q9AEN7	Q9AEN7 rhizobium s
41	45	41.3	344	2	Q92KY2	Q92KY2 rhizobium s
42	45	41.3	505	2	Q96I80	Q96I80 homo sapien
43	45	41.3	607	2	P73606	P73606 synechocyst
44	45	41.3	875	1	NETR_HUMAN	P56730 homo sapien
45	44	40.4	63	2	Q9BS83	Q9BS83 homo sapien
46	44	40.4	121	2	Q8Q0W8	Q8Q0W8 methanosarc
47	44	40.4	172	2	Q897T4	Q897T4 clostridium
48	44	40.4	338	2	Q8SSM8	Q8SSM8 encephalito
49	44	40.4	385	1	DIAC_HUMAN	Q01459 homo sapien
50	44	40.4	485	2	Q8CFK1	Q8CFK1 mus musculu
51	44	40.4	498	1	NMT2_HUMAN	O60551 mus sapien
52	44	40.4	529	1	NMT2_MOUSE	O70311 mus musculu
53	44	40.4	529	2	Q700Q7	Q700Q7 rattus norv
54	44	40.4	918	2	Q6BQV5	Q6BQV5 debaromyce
55	44	40.4	1193	2	Q7Z8L8	Q7Z8L8 cochliobolu
56	44	40.4	1907	2	Q96L96	Q96L96 homo sapien
57	43.5	39.9	339	2	Q90709	Q90709 cowpea aphl
58	43.5	39.9	476	2	Q6B8Q6	Q6B8Q6 gracilaria
59	43.5	39.9	635	2	Q8TUQ1	Q8TUQ1 methanosarc
60	43	39.4	37	2	Q8P0M2	Q8P0M2 streptococc
61	43	39.4	150	2	Q6TUT7	Q6TUT7 yaba monkey
62	43	39.4	151	2	Q8PRK4	Q8PRK4 xanthomonas
63	43	39.4	190	2	Q6K3W6	Q6K3W6 oryza sativ
64	43	39.4	206	2	Q8PKC5	Q8PKC5 xanthomonas
65	43	39.4	239	2	O44730	O44730 caenorhabdi
66	43	39.4	274	2	Q6AHQ6	Q6AHQ6 caenorhabdi
67	43	39.4	310	2	Q7RFC4	Q7RFC4 plasmodium
68	43	39.4	319	2	Q8UDV8	Q8UDV8 agrobacteri
69	43	39.4	341	1	Y5X3_CAEBL	Q10022 caenorhabdi
70	43	39.4	342	2	Q6AHQ7	Q6AHQ7 caenorhabdi
71	43	39.4	356	2	Q919T1	Q919T1 influenza a
72	43	39.4	363	2	Q82BN3	Q82BN3 streptomyce
73	43	39.4	369	2	Q7CY30	Q7CY30 agrobacteri
74	43	39.4	371	2	Q6KNR8	Q6KNR8 arabidopsis
75	43	39.4	394	2	Q8EVB5	Q8EVB5 mycoplasma
76	43	39.4	403	2	Q6MQE4	Q6MQE4 bdellovibri
77	43	39.4	428	2	Q8TBF1	Q8TBF1 homo sapien
78	43	39.4	489	2	Q8N9B8	Q8N9B8 homo sapien
79	43	39.4	505	1	SMI1_YEAST	P32566 saccharomyc
80	43	39.4	542	1	HNA2_CUCSA	P49295 cucumis sat
81	43	39.4	565	2	Q8NY08	Q8NY08 staphylococ
82	43	39.4	565	2	Q99WC5	Q99WC5 staphylococ
83	43	39.4	565	2	Q7A7C7	Q7A7C7 staphylococ
84	43	39.4	565	2	Q6GC10	Q6GC10 staphylococ
85	43	39.4	565	2	Q6GJJ5	Q6GJJ5 staphylococ
86	43	39.4	605	2	Q8PZF0	Q8PZF0 methanosarc
87	43	39.4	639	2	Q7QUK1	Q7QUK1 giardia lam
88	43	39.4	750	2	Q26188	Q26188 plasmodium
89	43	39.4	850	2	Q7RIR5	Q7RIR5 plasmodium
90	43	39.4	1150	2	P79045	P79045 emericella
91	43	39.4	1315	2	O812E4	O812E4 plasmodium
92	43	39.4	1322	2	Q6UDM4	Q6UDM4 plasmodium
93	43	39.4	1324	2	Q6UDW2	Q6UDW2 plasmodium
94	43	39.4	1325	2	Q6LFQ7	Q6LFQ7 plasmodium
95	43	39.4	1327	2	Q9NFB4	Q9NFB4 plasmodium
96	43	39.4	1778	2	Q9BPP7	Q9BPP7 antheraea y
97	43	39.4	1778	2	Q9GUX5	Q9GUX5 antheraea p
98	43	39.4	2171	1	CCAC_RABIT	P15381 o voltage-d
99	42	38.5	73	2	Q7QPI9	Q7QPI9 giardia lam
100	42	38.5	127	2	Q7YDJ6	Q7YDJ6 ranunculus
101	42	38.5	130	1	RS8_YERPE	Q8Z198 versinia pe
102	42	38.5	130	2	O664T5	O664T5 versinia pe
103	42	38.5	137	2	Q8NAF9	Q8NAF9 homo sapien
104	42	38.5	182	1	YLP3_CAEBL	P34383 caenorhabdi

105	42	38.5	226	2	Q69XV4	Q69XV4 oryza sativ	178	41	37.6	752	1	RRP1_INBP9	O36430 influenza b
106	42	38.5	257	2	Q6GR33	Q6gr33 xenopus lae	179	41	37.6	752	2	Q68IP2	Q68ip2 influenza b
107	42	38.5	265	2	Q86KQ7	Q86kq7 dictyosteli	180	41	37.6	752	2	Q77SM9	Q77sm9 influenza b
108	42	38.5	274	2	Q6GN60	Q6gn60 xenopus lae	181	41	37.6	752	2	Q77SN0	Q77sn0 influenza b
109	42	38.5	305	2	Q9RDP5	Q9rdp5 streptomyce	182	41	37.6	752	2	Q77SN1	Q77sn1 influenza b
110	42	38.5	356	2	Q9N3M8	Q9n3m8 caenorhabdi	183	41	37.6	752	2	Q77SN2	Q77sn2 influenza b
111	42	38.5	362	2	Q81IE3	Q81ie3 plasmodium	184	41	37.6	752	2	Q80DM8	Q80dm8 influenza b
112	42	38.5	375	2	Q96Y62	Q96y62 sulfolobus	185	41	37.6	752	2	Q80DN5	Q80dn5 influenza b
113	42	38.5	402	1	ISPG_SYNEL	Q8dk70 synechococc	186	41	37.6	752	2	Q9IMP8	Q9imp8 influenza b
114	42	38.5	457	2	Q750A7	Q750a7 ashbya goss	187	41	37.6	752	2	Q9PWU1	Q9pwu1 influenza b
115	42	38.5	461	2	Q69XV5	Q69xv5 oryza sativ	188	41	37.6	752	2	Q9QLJ8	Q9qlj8 influenza b
116	42	38.5	513	2	Q9A731	Q9a731 caulobacter	189	41	37.6	752	2	Q9QLJ9	Q9qlj9 influenza b
117	42	38.5	535	2	Q6BUX0	Q6bux0 debaryomyce	190	41	37.6	752	2	Q9QLK0	Q9qlk0 influenza b
118	42	38.5	574	2	Q8QOC1	Q8qoc1 pseudomonas	191	41	37.6	752	2	Q9QLK1	Q9qlk1 influenza b
119	42	38.5	731	2	Q86557	Q86557 streptomyce	192	41	37.6	752	2	Q9QLK2	Q9qlk2 influenza b
120	42	38.5	776	2	Q785X4	Q785x4 neurospora	193	41	37.6	752	2	Q9QLK3	Q9qlk3 influenza b
121	42	38.5	787	2	Q96020	Q96u20 neurospora	194	41	37.6	752	2	Q9QLK4	Q9qlk4 influenza b
122	42	38.5	845	2	Q8WQJ3	Q8wqj3 plasmodium	195	41	37.6	752	2	Q9QLK5	Q9qlk5 influenza b
123	42	38.5	915	2	Q7UV66	Q7uv66 rhodopirell	196	41	37.6	752	2	Q9QLK6	Q9qlk6 influenza b
124	42	38.5	957	2	Q9K760	Q9k760 bacillus ha	197	41	37.6	752	2	Q9QLK7	Q9qlk7 influenza b
125	42	38.5	973	2	Q9NC89	Q9nc89 strongyloce	198	41	37.6	752	2	Q9QLK8	Q9qlk8 influenza b
126	42	38.5	1075	2	Q9NC90	Q9nc90 strongyloce	199	41	37.6	752	2	Q9QLK9	Q9qlk9 influenza b
127	42	38.5	1304	2	Q8WT63	Q8wt63 plasmodium	200	41	37.6	888	2	Q75AE3	Q75ae3 ashbya goss
128	42	38.5	1319	2	Q78DE1	Q78del neurospora	201	41	37.6	1233	1	RGY2_SULTO	Q97322 plasmodium
129	42	38.5	1450	2	Q815X9	Q815x9 plasmodium	202	41	37.6	1334	2	Q97322	Q97322 sulfolobus
130	42	38.5	1450	2	Q815X9	Q815x9 plasmodium	202	41	37.6	1334	2	Q97322	Q97322 plasmodium
131	42	38.5	1919	2	Q8LRK9	Q8lrk9 arabidopsis	203	41	37.6	1969	2	Q69HL6	Q69hl6 ciona intes
132	42	38.5	1994	2	Q9LP19	Q9lp19 arabidopsis	204	40.5	37.2	205	2	Q6FZU4	Q6fzu4 bartonella
133	42	38.5	3115	2	Q9VQV1	Q9vqv1 drosophila	205	40.5	37.2	223	2	Q801A6	Q801a6 latimeria m
134	42	38.5	4280	2	Q9UB29	Q9ub29 caenorhabdi	206	40.5	37.2	247	2	Q7VN89	Q7vnr9 haemophilus
135	42	38.5	4450	2	Q9UB28	Q9ub28 caenorhabdi	207	40.5	37.2	252	2	Q9LLQ7	Q9llq7 streptomyce
136	41.5	38.1	19066	2	Q801W8	Q801w8 brachydanio	208	40.5	37.2	302	2	Q9HXC1	Q9hxc1 pseudomonas
137	41.5	38.1	325	2	Q9PP25	Q9pp25 oryza sativ	209	40.5	37.2	431	2	Q56454	Q56454 paracoccus
138	41.5	38.1	325	2	Q72KA9	Q72ka9 campylobact	210	40.5	37.2	503	2	Q6FZW1	Q6fzw1 bartonella
139	41.5	38.1	470	2	Q66IF4	Q66if4 thermus the	211	40.5	37.2	507	2	Q8YGW3	Q8ygw3 brucella me
140	41	37.6	46	2	Q647D0	Q647d0 xenopus tro	212	40.5	37.2	507	2	Q8G100	Q8g100 brucella su
141	41	37.6	76	2	Q8S3A8	Q8s3a8 bacillus me	213	40.5	37.2	946	1	IHB_DROME	Q81643 drosophila
142	41	37.6	100	2	Q647D0	Q647d0 thermoprote	214	40.5	37.2	1679	2	Q81LA0	Q81la0 plasmodium
143	41	37.6	129	2	Q74NG1	Q74ng1 nanoarchaeu	215	40.5	37.2	3164	2	Q813E6	Q813e6 plasmodium
144	41	37.6	129	2	Q6XYX4	Q6xyx4 spiropasma	216	40	36.7	55	2	Q8S2X3	Q8s2x3 zea mays (m
145	41	37.6	130	2	RS8_BUCAI	RS8buc1 xenopus lae	217	40	36.7	56	2	Q8LNE9	Q8lne9 oryza sativ
146	41	37.6	137	2	Q78ZB1	Q78zb1 xenopus lae	218	40	36.7	56	2	Q8RUB4	Q8rub4 zea mays (m
147	41	37.6	147	1	Y237_TROWT	P67323 tropheryma	219	40	36.7	67	2	Q82548	Q82548 influenza a
148	41	37.6	147	1	Y533_TROW8	P67324 tropheryma	220	40	36.7	78	2	Q82547	Q82547 influenza a
149	41	37.6	154	2	Q8BUS2	Q8bus2 mus musculu	221	40	36.7	79	2	Q6Z419	Q6z419 oryza sativ
150	41	37.6	158	2	Q72V51	Q72v51 xenopus lae	222	40	36.7	84	2	Q8H522	Q8h522 oryza sativ
151	41	37.6	160	1	RRAL_VIBPA	Q8red2 vibrio para	223	40	36.7	124	2	Q91TY9	Q91ty9 influenza a
152	41	37.6	164	2	Q7TV85	Q7tv85 prochloroco	224	40	36.7	129	2	Q91TZ1	Q91tz1 influenza a
153	41	37.6	184	2	Q728K5	Q728k5 desulfovibr	225	40	36.7	130	1	RS8_BUCAP	P59030 buchnera ap
154	41	37.6	191	2	Q9AY18	Q9ay18 funaria apo	226	40	36.7	131	2	Q7NQG6	Q7nqg6 chromobacte
155	41	37.6	193	2	Q9AY16	Q9ay16 entosthodon	227	40	36.7	131	2	Q7V7B7	Q7v7b7 bordetella
156	41	37.6	238	2	Q74551	Q74551 schizosacch	228	40	36.7	131	2	Q7W2E1	Q7w2e1 bordetella
157	41	37.6	255	2	Q8MI30	Q8mi30 equus cabal	229	40	36.7	131	2	Q7WRA9	Q7wra9 bordetella
158	41	37.6	255	2	Q8DTU7	Q8dtu7 streptococc	230	40	36.7	132	1	RS8_TREPA	Q82333 treponema p
159	41	37.6	281	1	DPHS_MOUSE	Q9cwg0 mus musculu	231	40	36.7	150	2	Q9KCP2	Q9kcp2 bacillus ha
160	41	37.6	281	2	Q6PAC5	Q6pac5 mus musculu	232	40	36.7	162	2	Q8HZN2	Q8hzn2 elephas max
161	41	37.6	288	2	Q81CQ3	Q81cq3 bacillus ce	233	40	36.7	162	2	Q48953	Q48953 arabidopsis
162	41	37.6	294	2	Q7RPL7	Q7rpl7 plasmodium	234	40	36.7	176	2	Q9HIV8	Q9hiv8 thermoplasma
163	41	37.6	308	2	Q6LRD5	Q6lrd5 photobacter	235	40	36.7	187	2	Q8HU07	Q8hu07 rhodobryum
164	41	37.6	312	2	Q6HI38	Q6hi38 bacillus th	236	40	36.7	188	2	Q6JKP4	Q6jkp4 influenza a
165	41	37.6	314	2	Q7RJ65	Q7rj65 plasmodium	237	40	36.7	188	2	Q6JKP5	Q6jkp5 influenza a
166	41	37.6	332	2	Q7QB96	Q7qb96 anopheles g	238	40	36.7	188	2	Q6JKP6	Q6jkp6 influenza a
167	41	37.6	373	2	Q6D5A6	Q6d5a6 erwinia car	239	40	36.7	188	2	Q6JKP7	Q6jkp7 influenza a
168	41	37.6	437	2	Q871H0	Q871h0 neurospora	240	40	36.7	191	2	Q5W760	Q5w760 influenza a
169	41	37.6	481	2	Q757L7	Q757l7 neurospora	241	40	36.7	191	2	Q5W761	Q5w761 influenza a
170	41	37.6	644	2	Q7RCW9	Q7rcw9 plasmodium	242	40	36.7	191	2	Q8BCD5	Q8bcd5 influenza a
171	41	37.6	721	2	Q9PA43	Q9pa43 xyella fas	243	40	36.7	191	2	Q8BCD6	Q8bcd6 influenza a
172	41	37.6	730	2	Q7SHT6	Q7sht6 xyella fas	244	40	36.7	191	2	Q8BCD7	Q8bcd7 influenza a
173	41	37.6	739	2	Q9Q0S6	Q9q0s6 influenza a	245	40	36.7	191	2	Q8BCD8	Q8bcd8 influenza a
174	41	37.6	747	2	Q8QPM0	Q8qpm0 influenza b	246	40	36.7	191	2	Q8BCD9	Q8bcd9 influenza a
175	41	37.6	752	1	RRP1_INBAC	P13871 influenza b	247	40	36.7	191	2	Q8BCE0	Q8bce0 influenza a
176	41	37.6	752	1	RRP1_INBAD	P13872 influenza b	248	40	36.7	191	2	Q8BCE1	Q8bce1 influenza a
177	41	37.6	752	1	RRP1_INBLE	P07832 influenza b	249	40	36.7	191	2	Q8BCE2	Q8bce2 influenza a
							250	40	36.7	191	2	Q8BCE3	Q8bce3 influenza a

251	40	36.7	191	2	Q8BCE4	Q8bce4 influenza a	324	40	36.7	647	2	Q809M6	Q809m6 influenza a
252	40	36.7	213	2	Q89S17	Q89s17 brachyrihizob	325	40	36.7	647	2	Q809M7	Q809m7 influenza a
253	40	36.7	252	2	Q04455	Q04455 pseudomonas	326	40	36.7	660	2	Q71FW4	Q71fw4 influenza a
254	40	36.7	256	2	Q8BC95	Q8bc95 influenza a	327	40	36.7	683	2	Q6DNU0	Q6dnu0 influenza a
255	40	36.7	280	2	Q98KM9	Q98km9 rhizobium l	328	40	36.7	683	2	Q71FV9	Q71fv9 influenza a
256	40	36.7	283	2	Q7R7D1	Q7r7d1 plasmodium	329	40	36.7	689	2	Q71FW5	Q71fw5 influenza a
257	40	36.7	288	2	Q9DW23	Q9dw23 rat cytomeg	330	40	36.7	690	2	Q6DNS7	Q6dns7 influenza a
258	40	36.7	290	2	Q8X681	Q8x681 escherichia	331	40	36.7	692	2	Q82572	Q82572 influenza a
259	40	36.7	302	2	Q96OH4	Q96oh4 drosophila	332	40	36.7	696	2	Q71FW2	Q71fw2 influenza a
260	40	36.7	304	1	HEM3_XANCP	Q8p536 xanthomonas	333	40	36.7	696	2	Q778V2	Q778v2 influenza a
261	40	36.7	312	1	Q81PV8	Q81pv8 bacillus an	334	40	36.7	703	2	Q71FW0	Q71fw0 influenza a
262	40	36.7	315	1	MHPR_ECOLI	P77569 escherichia	335	40	36.7	704	2	Q67EK9	Q67ek9 influenza a
263	40	36.7	315	2	Q7AH54	Q7ah54 escherichia	336	40	36.7	706	2	Q6DNT6	Q6dnt6 influenza a
264	40	36.7	330	2	Q85ZW5	Q85zw5 mus musculus	337	40	36.7	713	2	Q67EK7	Q67ek7 influenza a
265	40	36.7	338	2	Q7R9R7	Q7r9r7 plasmodium	338	40	36.7	717	2	Q67EJ8	Q67ej8 influenza a
266	40	36.7	362	2	Q61L91	Q61l91 drosophila	339	40	36.7	717	2	Q67EK8	Q67ek8 influenza a
267	40	36.7	363	2	Q673E6	Q673e6 medicago tr	340	40	36.7	717	2	Q67EL4	Q67el4 influenza a
268	40	36.7	368	2	Q7RSW6	Q7rsw6 plasmodium	341	40	36.7	718	2	Q67EK0	Q67ek0 influenza a
269	40	36.7	381	2	Q8QCL9	Q8ql9 rhizobium l	342	40	36.7	718	2	Q6DNS6	Q6dns6 influenza a
270	40	36.7	381	2	Q8QPH7	Q8qph7 influenza a	343	40	36.7	720	2	Q9PZF3	Q9pzf3 influenza a
271	40	36.7	382	2	Q9QOK4	Q9qok4 influenza a	344	40	36.7	721	2	Q6J850	Q6j850 influenza a
272	40	36.7	382	2	Q9QOK5	Q9qok5 influenza a	345	40	36.7	722	2	Q6DNS3	Q6dns3 influenza a
273	40	36.7	382	2	Q9QOK6	Q9qok6 influenza a	346	40	36.7	726	2	Q67EJ9	Q67ej9 influenza a
274	40	36.7	382	2	Q9QJC1	Q9qjc1 influenza a	347	40	36.7	726	2	Q67EK6	Q67ek6 influenza a
275	40	36.7	399	1	PNCB_ECO57	Q8xds8 escherichia	348	40	36.7	726	2	Q67EL3	Q67el3 influenza a
276	40	36.7	399	1	PNCB_ECOLI	P18133 escherichia	349	40	36.7	726	2	Q67EL5	Q67el5 influenza a
277	40	36.7	399	1	PNCB_SALTI	Q8z7y9 salmonella	350	40	36.7	727	2	Q6J848	Q6j848 influenza a
278	40	36.7	399	1	PNCB_SALTY	P22253 salmonella	351	40	36.7	728	2	Q71FW1	Q71fw1 influenza a
279	40	36.7	400	2	Q7UD27	Q7ud27 shigella fl	352	40	36.7	730	1	SMIA_SCHAM	Q26473 schistocerc
280	40	36.7	401	2	Q82N33	Q82n33 streptomyce	353	40	36.7	736	2	Q6J855	Q6j855 influenza a
281	40	36.7	416	2	Q83LN3	Q83ln3 shigella fl	354	40	36.7	736	2	Q9QOS2	Q9qos2 influenza a
282	40	36.7	416	2	Q8FJ98	Q8fj98 escherichia	355	40	36.7	738	2	Q9PY91	Q9py91 influenza a
283	40	36.7	422	1	PGLI_ARATH	P49062 arabidopsis	356	40	36.7	739	2	Q71FW6	Q71fw6 influenza a
284	40	36.7	422	1	Q9K426	Q9k426 streptomyce	357	40	36.7	739	2	Q9QOS5	Q9qos5 influenza a
285	40	36.7	430	2	Q67EK3	Q67ek3 influenza a	358	40	36.7	742	2	Q9PZH8	Q9pzh8 influenza a
286	40	36.7	430	2	Q67EK4	Q67ek4 influenza a	359	40	36.7	742	2	Q9PZH9	Q9pzh9 influenza a
287	40	36.7	430	2	Q67EK5	Q67ek5 influenza a	360	40	36.7	742	2	Q9PZ10	Q9pzi0 influenza a
288	40	36.7	435	2	O16968	O16968 caenorhabdi	361	40	36.7	742	2	Q9PZ11	Q9pzi1 influenza a
289	40	36.7	445	2	Q9GLJ39	Q9glj39 bos taurus	362	40	36.7	742	2	Q9PZ12	Q9pzi2 influenza a
290	40	36.7	463	1	Q67EK1	Q67ek1 influenza a	363	40	36.7	742	2	Q9PZ13	Q9pzi3 influenza a
291	40	36.7	465	1	HMA2_HORVU	Q96563 hordeum vul	364	40	36.7	742	2	Q9PZ15	Q9pzi5 influenza a
292	40	36.7	466	2	Q67EK2	Q67ek2 influenza a	365	40	36.7	742	2	Q9PZ19	Q9pzi9 influenza a
293	40	36.7	470	2	Q67EL2	Q67el2 influenza a	366	40	36.7	742	2	Q9QOS7	Q9qos7 influenza a
294	40	36.7	471	1	YIRI_DROME	P16424 drosophila	367	40	36.7	744	2	Q71FW3	Q71fw3 influenza a
295	40	36.7	479	2	Q919V1	Q919v1 influenza a	368	40	36.7	744	2	Q9QOS4	Q9qos4 influenza a
296	40	36.7	479	2	Q919V2	Q919v2 influenza a	369	40	36.7	745	2	Q6W8X6	Q6w8x6 influenza a
297	40	36.7	479	2	Q919V3	Q919v3 influenza a	370	40	36.7	745	2	Q9QOR8	Q9qor8 influenza a
298	40	36.7	488	2	Q8XT25	Q8xt25 ralestonia s	371	40	36.7	745	2	Q9QOS3	Q9qos3 influenza a
299	40	36.7	482	2	Q8QPH6	Q8qph6 influenza a	372	40	36.7	746	2	Q8QPH8	Q8qph8 influenza a
300	40	36.7	496	2	Q8QPH5	Q8qph5 influenza a	373	40	36.7	746	2	Q9QOR9	Q9qor9 influenza a
301	40	36.7	497	2	Q75KY5	Q75ky5 oryza sativ	374	40	36.7	746	2	Q9QOS0	Q9qos0 influenza a
302	40	36.7	498	1	NMT2_BOVIN	Q9n181 bos taurus	375	40	36.7	747	2	Q6DNT7	Q6dnt7 influenza a
303	40	36.7	508	2	Q8UY48	Q8uy48 influenza a	376	40	36.7	747	2	Q6DNT1	Q6dnt1 influenza a
304	40	36.7	525	2	Q6FS20	Q6fa20 candida gla	377	40	36.7	748	2	Q9QOS8	Q9qos8 influenza a
305	40	36.7	527	1	HMA1_HORVU	Q42843 hordeum vul	378	40	36.7	749	2	Q9QOH3	Q9qoh3 influenza a
306	40	36.7	528	2	Q97379	Q97379 strongyloce	379	40	36.7	750	2	Q8UY47	Q8uy47 influenza a
307	40	36.7	530	1	HMA2_ARATH	P49294 arabidopsis	380	40	36.7	751	2	Q8QM22	Q8qm22 influenza a
308	40	36.7	536	1	HEM1_ORYSA	O48674 oryza sativ	381	40	36.7	751	2	Q8QPH3	Q8qph3 influenza a
309	40	36.7	536	2	Q9FW00	Q9fw00 oryza sativ	382	40	36.7	751	2	Q8QPH4	Q8qph4 influenza a
310	40	36.7	537	2	Q9V7E0	Q9v7e0 drosophila	383	40	36.7	751	2	Q8QPH9	Q8qph9 influenza a
311	40	36.7	553	2	Q6ASN7	Q6asm7 propionibac	384	40	36.7	751	2	Q91CC5	Q91cc5 influenza a
312	40	36.7	577	1	YQ4B_CAEEL	Q17426 caenorhabdi	385	40	36.7	751	2	Q91CC6	Q91cc6 influenza a
313	40	36.7	586	2	Q62YB1	Q62yb1 bacillus li	386	40	36.7	751	2	Q71FW0	Q71fw0 influenza a
314	40	36.7	586	2	Q71FU9	Q71fu9 influenza a	387	40	36.7	752	2	Q8JN15	Q8jnl5 influenza a
315	40	36.7	587	1	HE_PARLI	P22757 paracentrot	388	40	36.7	752	2	Q71FW6	Q71fw6 influenza a
316	40	36.7	587	2	Q65MW8	Q65mw8 bacillus li	389	40	36.7	752	2	Q71FW8	Q71fw8 influenza a
317	40	36.7	588	2	Q70681	Q70681 sugarcane s	390	40	36.7	752	2	Q71FW3	Q71fw3 influenza a
318	40	36.7	619	2	Q6E1Y9	Q6eiy9 canis faml	391	40	36.7	755	2	Q8JT71	Q8jtc71 influenza a
319	40	36.7	628	2	Q6DNS4	Q6dns4 influenza a	392	40	36.7	755	2	Q8QM32	Q8qm32 influenza a
320	40	36.7	628	2	Q6J845	Q6j845 influenza a	393	40	36.7	755	2	Q6DTW6	Q6dtw6 influenza a
321	40	36.7	633	2	Q8MLA8	Q8mla8 drosophila	394	40	36.7	755	2	Q6DXT4	Q6dtx4 influenza a
322	40	36.7	633	2	Q7PIA3	Q7pia3 chromobacte	395	40	36.7	756	2	Q8QM23	Q8qm23 influenza a
323	40	36.7	644	1	SYR_AERPE	Q9yb39 aeropyrum p	396	40	36.7	756	2	Q8QM25	Q8qm25 influenza a

397	40	36.7	756	2	Q8QM26	Q8qm26	influenza a
398	40	36.7	756	2	Q8QM28	Q8qm28	influenza a
399	40	36.7	756	2	Q91E77	Q91e77	influenza a
400	40	36.7	756	2	Q6DNQ8	Q6dnq8	influenza a
401	40	36.7	756	2	Q6DNS8	Q6dns8	influenza a
402	40	36.7	756	2	Q6DNT0	Q6dnt0	influenza a
403	40	36.7	756	2	Q6DNT1	Q6dnt1	influenza a
404	40	36.7	756	2	Q6DNT2	Q6dnt2	influenza a
405	40	36.7	756	2	Q6DNT3	Q6dnt3	influenza a
406	40	36.7	756	2	Q6DNT4	Q6dnt4	influenza a
407	40	36.7	756	2	Q6DNT5	Q6dnt5	influenza a
408	40	36.7	756	2	Q6DNT8	Q6dnt8	influenza a
409	40	36.7	756	2	Q6DNT9	Q6dnt9	influenza a
410	40	36.7	756	2	Q6DNU1	Q6dnu1	influenza a
411	40	36.7	756	2	Q6DNU2	Q6dnu2	influenza a
412	40	36.7	756	2	Q6DNU3	Q6dnu3	influenza a
413	40	36.7	756	2	Q6DNU9	Q6dnu9	influenza a
414	40	36.7	756	2	Q6DNV0	Q6dnv0	influenza a
415	40	36.7	756	2	Q6DNV1	Q6dnv1	influenza a
416	40	36.7	757	1	RRP1_TAANN	P21426	influenza a
417	40	36.7	757	1	RRP1_IABEI	P16502	influenza a
418	40	36.7	757	1	RRP1_IABOM	Q82571	influenza a
419	40	36.7	757	1	RRP1_IAGU2	P16503	influenza a
420	40	36.7	757	1	RRP1_IATHE	P16504	influenza a
421	40	36.7	757	1	RRP1_IATHE	P16505	influenza a
422	40	36.7	757	1	RRP1_IAKIE	P18882	influenza a
423	40	36.7	757	1	RRP1_IAKIT	Q91741	influenza a
424	40	36.7	757	1	RRP1_IAKOR	P16506	influenza a
425	40	36.7	757	1	RRP1_IALF1	P26119	influenza a
426	40	36.7	757	1	RRP1_IALF2	P26120	influenza a
427	40	36.7	757	1	RRP1_IALF3	P26121	influenza a
428	40	36.7	757	1	RRP1_IAMAN	P16508	influenza a
429	40	36.7	757	1	RRP1_IAMT8	P16507	influenza a
430	40	36.7	757	1	RRP1_IANT6	P03432	influenza a
431	40	36.7	757	1	RRP1_IAPUE	P03431	influenza a
432	40	36.7	757	1	RRP1_IASIN	P16511	influenza a
433	40	36.7	757	1	RRP1_IATKM	P16513	influenza a
434	40	36.7	757	1	RRP1_IAVI7	P31341	influenza a
435	40	36.7	757	1	RRP1_IAWIL	P03430	influenza a
436	40	36.7	757	1	RRP1_IAWIS	P16514	influenza a
437	40	36.7	757	1	RRP1_IAWI3	P16509	influenza a
438	40	36.7	757	1	RRP1_IAZON	P16510	influenza a
439	40	36.7	757	1	RRP1_IATZF	P16512	influenza a
440	40	36.7	757	2	Q8JUD8	Q8jld8	influenza a
441	40	36.7	757	2	Q8JSN8	Q8jsn8	influenza a
442	40	36.7	757	2	Q8JSY7	Q8jsy7	influenza a
443	40	36.7	757	2	Q8JT64	Q8jt64	influenza a
444	40	36.7	757	2	Q8JT82	Q8jt82	influenza a
445	40	36.7	757	2	Q8JU07	Q8juu7	influenza a
446	40	36.7	757	2	Q8QT42	Q8qt42	influenza a
447	40	36.7	757	2	Q8QT43	Q8qt43	influenza a
448	40	36.7	757	2	Q8QT44	Q8qt44	influenza a
449	40	36.7	757	2	Q8QT45	Q8qt45	influenza a
450	40	36.7	757	2	Q8QT46	Q8qt46	influenza a
451	40	36.7	757	2	Q8QT47	Q8qt47	influenza a
452	40	36.7	757	2	Q8QT48	Q8qt48	influenza a
453	40	36.7	757	2	Q8QT49	Q8qt49	influenza a
454	40	36.7	757	2	Q8V3X6	Q8v3x6	influenza a
455	40	36.7	757	2	Q8V3Y2	Q8v3y2	influenza a
456	40	36.7	757	2	Q910D6	Q910d6	influenza a
457	40	36.7	757	2	Q91HA4	Q91ha4	influenza a
458	40	36.7	757	2	Q98VV0	Q98vv0	influenza a
459	40	36.7	757	2	Q98VV2	Q98vv2	influenza a
460	40	36.7	757	2	Q995P1	Q995p1	influenza a
461	40	36.7	757	2	Q995P2	Q995p2	influenza a
462	40	36.7	757	2	Q66WML	Q66wml	influenza a
463	40	36.7	757	2	Q67322	Q67322	influenza a
464	40	36.7	757	2	Q692M3	Q692m3	influenza a
465	40	36.7	757	2	Q693C0	Q693c0	influenza a
466	40	36.7	757	2	Q693Q0	Q693q0	influenza a
467	40	36.7	757	2	Q6DNC3	Q6dnc3	influenza a
468	40	36.7	757	2	Q6DNS9	Q6dns9	influenza a
469	40	36.7	757	2	Q6DSU8	Q6dsu8	influenza a
470	40	36.7	757	2	Q6E3N5	Q6e3n5	influenza a
471	40	36.7	757	2	Q6E3N6	Q6e3n6	influenza a
472	40	36.7	757	2	Q6E3N7	Q6e3n7	influenza a
473	40	36.7	757	2	Q6E3N8	Q6e3n8	influenza a
474	40	36.7	757	2	Q6E3N9	Q6e3n9	influenza a
475	40	36.7	757	2	Q6E3P1	Q6e3p1	influenza a
476	40	36.7	757	2	Q6E3P2	Q6e3p2	influenza a
477	40	36.7	757	2	Q6E3P3	Q6e3p3	influenza a
478	40	36.7	757	2	Q6E3P4	Q6e3p4	influenza a
479	40	36.7	757	2	Q6E3P5	Q6e3p5	influenza a
480	40	36.7	757	2	Q6E3P6	Q6e3p6	influenza a
481	40	36.7	757	2	Q6E3P7	Q6e3p7	influenza a
482	40	36.7	757	2	Q6E3P8	Q6e3p8	influenza a
483	40	36.7	757	2	Q6E3P9	Q6e3p9	influenza a
484	40	36.7	757	2	Q6E3Q0	Q6e3q0	influenza a
485	40	36.7	757	2	Q6E3Q2	Q6e3q2	influenza a
486	40	36.7	757	2	Q6E3Q3	Q6e3q3	influenza a
487	40	36.7	757	2	Q6E3Q4	Q6e3q4	influenza a
488	40	36.7	757	2	Q6E3Q5	Q6e3q5	influenza a
489	40	36.7	757	2	Q61WJ3	Q61wj3	influenza a
490	40	36.7	757	2	Q6J0Q4	Q6j0q4	influenza a
491	40	36.7	757	2	Q6J851	Q6j851	influenza a
492	40	36.7	757	2	Q6J856	Q6j856	influenza a
493	40	36.7	757	2	Q6PNE2	Q6pne2	influenza a
494	40	36.7	757	2	Q6R2G3	Q6r2g3	influenza a
495	40	36.7	757	2	Q6VGI1	Q6vgi1	influenza a
496	40	36.7	757	2	Q6VGI2	Q6vgi2	influenza a
497	40	36.7	757	2	Q6VGI3	Q6vgi3	influenza a
498	40	36.7	757	2	Q6W8X7	Q6w8x7	influenza a
499	40	36.7	757	2	Q6XS26	Q6xs26	influenza a
500	40	36.7	757	2	Q6XS27	Q6xs27	influenza a

ALIGNMENTS

RESULT 1
LPP3_HUMAN
ID_LPP3_HUMAN STANDARD; PRT; 311 AA.
AC O14495; Q96GWD; Q99782;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Lipid phosphate phosphohydrolase 3 (EC 3.1.3.4) (Phosphatidic acid
phosphatase 2b) (Phosphatidate phosphohydrolase type 2b) (PAP2b) (PAP-
2b) (PAP2-beta) (Vascular endothelial growth factor and type I
collagen inducible protein) (VCIP).
GN Name=PPAP2b; Synonyms=LPP3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., CHARACTERIZATION, AND INDUCTION.
RX MEDLINE=97450990; PubMed=9305923; DOI=10.1074/jbc.272.39.24572;
RA Kai M., Wada I., Imai S.-I., Sakane F., Kanch H.;
RT "Cloning and characterization of two human isozymes of Mg2+-
independent phosphatidic acid phosphatase.";
RL J. Biol. Chem. 272:24572-24578(1997).
RN [2]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=98371049; PubMed=9705349; DOI=10.1074/jbc.273.34.22059;
RA Roberts R., Sciorra V.A., Morris A.J.;
RT "Human type 2 phosphatidic acid phosphohydrolases. Substrate
specificity of the type 2a, 2b, and 2c enzymes and cell surface
activity of the 2a isoform.";
RL J. Biol. Chem. 273:22059-22067(1998).
RN [3]
RP SEQUENCE FROM N.A.
RX PubMed=12660161; DOI=10.1093/emboj/cdg165;
RA Huntsree J.O., Feng S., Thakker G.D., Yang J., Hong J., Wary K.K.;
RT "Regulation of cell-cell interactions by phosphatidic acid phosphatase
2b/VCIP.";

EMBO J. 22:1539-1554(2003).
 [4]
 CC SEQUENCE FROM N.A.
 RN Leung D.W., Tompkins C.K.;
 RA "Molecular cloning and expression of an isoform of human
 RT phosphatidic acid phosphatase cDNA";
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 [5]
 CC SEQUENCE FROM N.A.
 RN MEDLINE=97264341; PubMed=9110174;
 RX Yu W., Andersson B., Worley K.C., Muzny D.M., Ding Y., Liu W.,
 RA Ricafrente J.Y., Wentland M.A., Lennon G., Gibbs R.A.;
 RT "Large-scale concatenation cDNA sequencing";
 RL Genome Res. 7:353-358(1997).
 [6]
 CC SEQUENCE FROM N.A.
 RN TISSUE=Placenta;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueclin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullighy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kertteman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RL and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 [7]
 CC SUBUNIT.
 RX PubMed=14725715; DOI=10.1186/1471-2091-5-2;
 RA Burnett C., Makridou P., Hewlett L., Howard K.;
 RT "Lipid phosphate phosphatases dimerise, but this interaction is not
 RL required for in vivo activity";
 BMC Biochem. 5:2-2(2004).
 CC -I- FUNCTION: Catalyzes the conversion of phosphatidic acid (PA) to
 CC diacylglycerol (DG). In addition it hydrolyzes lysophosphatidic
 CC acid (LPA) ceramide-1-phosphate (C-1-P) and sphingosine-1-
 CC phosphate (S-1-P). The relative catalytic efficiency is LPA = PA >
 CC C-1-P > S-1-P. May be involved in cell adhesion and in cell-cell
 CC interactions.
 CC -I- CATALYTIC ACTIVITY: A 3-sn-phosphatidate + H(2O) = a 1,2-diacyl-
 CC sn-glycerol + phosphate.
 CC -I- ENZYME REGULATION: Inhibited by sphingosine, zinc ions and
 CC propanolol. Not inhibited by N-ethylmaleimide treatment.
 CC -I- SUBUNIT: Homodimer. This complex seems not to be involved in
 CC substrate recognition, it may confer only structural or functional
 CC stability.
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein. Post-Golgi and
 CC plasma membrane localization.
 CC -I- TISSUE SPECIFICITY: Ubiquitously expressed. Highly expressed in
 CC heart and placenta.
 CC -I- INDUCTION: By epidermal growth factor (EGF), vascular endothelial
 CC growth factor (VEGF), basic fibroblast growth factor (bFGF) and
 CC phorbol myristate acetate (PMA).
 CC -I- PTM: N-glycosylated. Contains high-mannose oligosaccharides.
 CC -I- SIMILARITY: Belongs to the PA-phosphatase related phosphoesterase
 CC family.
 CC -I- CAUTION: Ref.5 sequence differs from that shown due to a
 CC frameshift in position 225.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 CC EMBL; AB000889; BAA22594.1; -;
 DR EMBL; AF017786; AAC63433.1; -;
 DR EMBL; AF480883; AAO84481.1; -;
 DR EMBL; AF043329; RAD02271.1; -;
 DR EMBL; U79294; AAB50222.1; ALT_FRAME.
 DR EMBL; BC003196; AAH03196.1; -;
 DR Genew; HGNC:9229; PPAP2B.
 DR H-InvDB; HIX0000628; -;
 DR Reactome; O14495; -;
 DR MIM; 607125; -;
 DR GO; GO:0016020; C:membrane; TAS.
 DR GO; GO:0004721; F:phosphoprotein phosphatase activity; TAS.
 DR GO; GO:0008151; P:cell growth and/or maintenance; TAS.
 DR GO; GO:0008354; P:germ-cell migration; TAS.
 DR InterPro; IPR008934; ACPase VanPerase.
 DR InterPro; IPR000326; Pesterase_PA_PTP.
 DR Pfam; PF01569; PAP2; 1.
 DR SMART; SM00014; acidppc; 1.
 KW Glycoprotein; Hydrolase; Transmembrane.
 FT DOMAIN 1 33 Cytoplasmic (Potential).
 FT TRANSMEM 34 54 Potential.
 FT DOMAIN 55 85 Lumenal (Potential).
 FT TRANSMEM 86 106 Potential.
 FT DOMAIN 107 122 Cytoplasmic (Potential).
 FT TRANSMEM 123 143 Potential.
 FT DOMAIN 144 193 Lumenal (Potential).
 FT TRANSMEM 194 214 Potential.
 FT DOMAIN 215 227 Cytoplasmic (Potential).
 FT TRANSMEM 228 248 Potential.
 FT DOMAIN 249 257 Lumenal (Potential).
 FT TRANSMEM 258 278 Potential.
 FT DOMAIN 279 311 Cytoplasmic (Potential).
 FT CARBOHYD 170 170 N-linked (GlcNAc...) (Potential).
 FT CONFLICT 282 282 T -> M (in Ref. 6).
 SQ SEQUENCE 311 AA; 35116 MW; CB3F60189044DA31 CRC64;
 Query Match 100.0%; Score 109; DB 1; Length 311;
 Best Local Similarity 100.0%; Pred. No. 9.9e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EGYIQRGDDSKVQEAR 20
 DB 173 EGYIQRGDDSKVQEAR 192
 RESULT 2
 LPP3 MOUSE
 ID LPP3 MOUSE STANDARD; PRT; 312 AA.
 AC Q99JY8; QB8TB7;
 DT 29-MAR-2004 (Rel. 43, Created)
 DT 29-MAR-2004 (Rel. 43, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Lipid phosphate phosphohydrolase 3 (EC 3.1.3.4) (Phosphatidic acid
 DE phosphatase 2b) (Phosphatidate phosphohydrolase type 2b) (PAP2b) (PAP-
 DE 2b) (PAP2-beta).
 GN Name=Ppap2b; Synonyms=Lpp3;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=C57BL/6J; TISSUE=Embryo;
 RC MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaido I., Osato N., Saito R., Saito H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,

CC	of axial structures. Loss of Ppap2b results in a marked increase
CC	in beta-catenin-mediated T-cell factor (TCF) transcription.
CC	-!- SIMILARITY: Belongs to the PA-phosphatase related phosphoesterase
CC	family.
CC	-!- CAUTION: Ref.1 sequence differs from that shown due to a
CC	frameshift in position 226.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; AK011276; BAC25327.1; ALT_FRAME.
DR	EMBL; BC005558; AAH05558.1; --.
DR	MGI; MGI:1915166; Ppap2b.
DR	GO; GO:0042577; F:lipid phosphatase activity; IMP.
DR	GO; GO:0001568; P:blood vessel development; IMP.
DR	GO; GO:0010003; P:gastulation (sensu Mammalia); IMP.
DR	GO; GO:0006644; P:phospholipid metabolism; IMP.
DR	GO; GO:0030111; P:regulation of wnt receptor signaling pathway; IDA.
DR	InterPro; IPR008934; Acpase Vanperase.
DR	Pfam; PF01569; PAP2; 1.
DR	SMART; SM00014; acidppc; 1.
KW	Developmental protein; Glycoprotein; Hydrolase; Transmembrane.
FT	DOMAIN 1 33 Cytoplasmic (Potential).
FT	TRANSMEM 34 54 Potential.
FT	DOMAIN 55 85 Lumenal (Potential).
FT	TRANSMEM 86 106 Potential.
FT	DOMAIN 107 123 Cytoplasmic (Potential).
FT	TRANSMEM 124 144 Potential.
FT	DOMAIN 145 184 Lumenal (Potential).
FT	TRANSMEM 195 215 Potential.
FT	DOMAIN 216 226 Cytoplasmic (Potential).
FT	TRANSMEM 227 247 Potential.
FT	DOMAIN 248 258 Lumenal (Potential).
FT	TRANSMEM 259 279 Potential.
FT	DOMAIN 280 312 Cytoplasmic (Potential).
FT	CARBOHYD 171 171 N-linked (GlcNAc...) (Potential).
SQ	SEQUENCE 312 AA; 35216 MW; D782986E04B57D7D CRC64;
Query Match 96.3%; Score 105; DB 1; Length 312;	
Best Local Similarity 95.0%; Pred.No. 4.7e-09;	
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;	
QY	1 EGYQNYRCRGDDSKVQEAR 20
DB	174 EGYQNYRCRGEDSKVQEAR 193
RESULT 3	
LPP3 RAT	
ID	LPP3 RAT STANDARD; PRT; 312 AA.
AC	P9754f.
DT	29-MAR-2004 (Rel. 43, Created)
DT	29-MAR-2004 (Rel. 43, Last sequence update)
DE	05-JUL-2004 (Rel. 44, Last annotation update)
DE	Lipid phosphate phosphohydrolase 3 (EC 3.1.3.4) (Phosphatidic acid
DE	phosphatase 2b) (Phosphatidate phosphohydrolase type 2b) (PAP2b) (PAP-
DE	2b) (PAP2-beta) (Differentially expressed in rat intestine 42)
DE	(Dr142).
GN	Name=Ppap2b; Synonyms=Lpp3;
OS	Rattus norvegicus (Rat).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX	NCBI_TaxID=10116;
RN	[1]
RP	SEQUENCE FROM N.A., TOPOLOGY, AND N-GLYCOSYLATION.
STRAIN=	Wistar; TISSU=Small intestine;
RC	MEDLINE=97094703; PubMed=8939937; DOI=10.1074/jbc.271.47.29928;

RA Barila D., Plateroti M., Nobili F., Muda A.O., Xie Y., Morimoto T.,
RA Perozzi G.;
RT "The Dri 42 gene, whose expression is upregulated during epithelial
RT differentiation, encodes a novel ER resident transmembrane protein.";
RL J. Biol. Chem. 271:29928-29936(1996).
RN [2]
RN TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
RX PubMed=8055940;
RX Barila D., Murgia C., Nobili F., Gaetani S., Perozzi G.;
RT "Subtractive hybridization cloning of novel genes differentially
RT expressed during intestinal development.";
RL Eur. J. Biochem. 223:701-709(1994).
RN [3]
RN TISSUE SPECIFICITY.
RP STRAIN=Sprague-Dawley;
RC MEDLINE=2155999; PubMed=11704545;
RX Nanjundan M., Possmayer F.;
RT "Molecular cloning and expression of pulmonary lipid phosphate
RT phosphohydrolases";
RL Am. J. Physiol. 281:L1484-L1493(2001).
CC -I- FUNCTION: Catalyzes the conversion of phosphatidic acid (PA) to
CC diacylglycerol (DG). In addition it hydrolyzes lysophosphatidic
CC acid (LPA), ceramide-1-phosphate (C-1-P) and sphingosine-1-
CC phosphate (S-1-P) (By similarity). Involved in the regulation of
CC epithelial differentiation.
CC -I- CATALYTIC ACTIVITY: A 3-sn-phosphatidate + H(2)O = a 1,2-diacyl-
CC sn-glycerol + phosphate.
CC -I- SUBUNIT: Homodimer (By similarity).
CC -I- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC reticulum associated.
CC -I- TISSUE SPECIFICITY: Detected in epithelial cells of intestinal
CC mucosa, lung, liver and brain.
CC -I- DEVELOPMENTAL STAGE: Expression is increased during epithelial
CC differentiation in intestinal mucosa as well as in kidney, liver
CC and lung.
CC -I- PTM: N-glycosylated.
CC -I- SIMILARITY: Belongs to the PA-phosphatase related phosphoesterase
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: Y07783; CAA69106.1; --
DR InterPro; IPR008934; AcPase_VanPerase.
DR InterPro; IPR000326; Pesterase_PA_PTP.
DR Pfam; PF01569; PAP2; 1.
DR SMART; SM00014; acidppc; 1.
KW Endoplasmic reticulum; Glycoprotein; Hydrolase; Transmembrane.
FT DOMAIN 1 33 Cytoplasmic (Potential).
FT TRANSMEM 34 54 Potential.
FT DOMAIN 55 85 Luminal (Potential).
FT TRANSMEM 86 106 Potential.
FT DOMAIN 107 123 Cytoplasmic (Potential).
FT TRANSMEM 124 144 Potential.
FT DOMAIN 145 194 Luminal (Potential).
FT TRANSMEM 195 215 Potential.
FT DOMAIN 216 226 Cytoplasmic (Potential).
FT TRANSMEM 227 247 Potential.
FT DOMAIN 248 258 Luminal (Potential).
FT TRANSMEM 259 279 Potential.
FT DOMAIN 280 312 Cytoplasmic (Potential).
FT CARBOHYD 171 171 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 312 AA; 35318 MW; 9B47FD321DB0419 CRC64;
Query Match 96.3%; Score 105; DB 1; Length 312;
Best Local Similarity 95.0%; Pred. No. 4.7e-09;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGYIQRRCRGDDSKVQEAR 20
DB 174 EGYIQRRCRGDDSKVQEAR 193
|||||
RESULT 4
Q6IMX4 PRELIMINARY; PRT; 312 AA.
ID Q6IMX4;
AC Q6IMX4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE ER transmembrane protein Dri 42.
GN Name=Dri42;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Heart;
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins E.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RN SEQUENCE FROM N.A.
RP TISSUE=Heart;
RC Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC072544; AAH72544.1; --
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR008934; AcPase_VanPerase.
DR InterPro; IPR000326; Pesterase_PA_PTP.
DR Pfam; PF01569; PAP2; 1.
DR SMART; SM00014; acidppc; 1.
KW Transmembrane.
SQ SEQUENCE 312 AA; 35233 MW; CDA54495C0E7D37D CRC64;
Query Match 96.3%; Score 105; DB 2; Length 312;
Best Local Similarity 95.0%; Pred. No. 4.7e-09;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGYIQRRCRGDDSKVQEAR 20
DB 174 EGYIQRRCRGDDSKVQEAR 193
|||||
RESULT 5
Q68F34 PRELIMINARY; PRT; 307 AA.
ID Q68F34;
AC Q68F34;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE MGC81884 protein.

GN Names=MGC81884;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1] SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatlenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
Whiting M., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
Krzyszynski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Klein S., Gerhard D.S.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC080011; RAH80011.1; -;
DR InterPro; IPR008934; AcPase_VanPerase.
DR InterPro; IPR000326; Pesterase_PA_PTP.
DR Pfam; PF01569; PAP2; 1.
DR SMART; SM00014; acidPPC; 1.
SQ SEQUENCE 307 AA; 34498 MW; 09666E6DA6265308 CRC64;
Query Match 67.9%; Score 74; DB 2; Length 307;
Best Local Similarity 68.4%; Pred. No. 0.00073;
Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 2 GYIYVRCRGDDSKVQEAR 20
Db |||:|||||:|||||
175 GYIYVRCRGPPNKVWEAR 193
ID Q6AX87 PRELIMINARY; PRT; 307 AA.
AC Q6AX87;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE MGC81990 protein.
OS Name=MGC81990;
GN Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.

OX NCBI_TaxID=8355;
RN [1] SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatlenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
Whiting M., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
Krzyszynski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Klein S., Gerhard D.S.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC079709; AAH79709.1; -;
DR InterPro; IPR008934; AcPase_VanPerase.
DR InterPro; IPR000326; Pesterase_PA_PTP.
DR Pfam; PF01569; PAP2; 1.
DR SMART; SM00014; acidPPC; 1.
SQ SEQUENCE 307 AA; 34417 MW; 23P956E7B8FFED0D CRC64;
Query Match 65.1%; Score 71; DB 2; Length 307;
Best Local Similarity 68.4%; Pred. No. 0.0023;
Matches 13; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 2 GYIYVRCRGDDSKVQEAR 20
Db |||:|||||:|||||
175 GYIYVRCRGPPNKVWEAR 193
ID Q95L84 PRELIMINARY; PRT; 167 AA.
AC Q95L84;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Phosphatidic acid phosphatase type 2A (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22306422; PubMed=12388084;
RA Wang D.-A., Du H., Jaggar J.H., Brindley D.N., Tigyi G.J.,
Watsky M.A.;
RT "Injury-elicited differential transcriptional regulation of

```

RT phospholipid growth factor receptors in the cornea.";
RL Am. J. Physiol. Cell Physiol. 283:C1646-C1654 (2002).
DR EMBL; AF04277; AAL01884.1; -.
DR InterPro; IPR008934; AcPase_VanPerase.
DR Pfam; PF01569; PAP2; 1.
DR SMART; SM00014; acidPPC; 1.
FT NON_TER 167
FT NON_TER 167
SQ SEQUENCE 167 AA; 18922 MW; 5185AECBC08CE9FB CRC64;

Query Match 64.2%; Score 70; DB 2; Length 167;
Best Local Similarity 60.0%; Pred. No. 0.0018; 4; Indels 0; Gaps 0;
Matches 12; Conservative 4; Mismatches PA;

QY 1 EGYIYNYRCRGDDSKVQEAR 20
DB 122 DGYIENYVCKGNAQKVEGR 141

RESULT 8
LPP1_RAT
ID LPP1_RAT STANDARD; PRT; 282 AA.
AC O08564; Q8K594;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Lipid phosphate phosphohydrolase 1 (EC 3.1.3.4) (Phosphatidic acid
DE phosphatase 2a) (Phosphatidate phosphohydrolase type 2a) (PAP2a) (PAP-
DE 2a) (PAP2-alpha).
CN Name=Pap2a; Synonyms=Lppl;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND CHARACTERIZATION.
RC Tissue=Liver;
RX PubMed=10359651; DOI=10.1042/0264-6021:3400677;
RA Jasinska R., Zhang Q.-X., Pilgill C., Singh I., Xu J., Dewald J.,
RA Dillon D.A., Berthiaume L.G., Carman G.M., Waggoner D.W.,
RA Brindley D.N.;
RT "Lipid phosphate phosphohydrolase-1 degrades exogenous glycerolipid
RT and sphingolipid phosphate esters.";
RL Biochem. J. 340:677-686 (1999).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC STRAIN=Sprague-Dawley; Tissue=Lung;
RX MEDLINE=21559999; PubMed=11704545;
RA Nanjundan M., Possemayer F.;
RT "Molecular cloning and expression of pulmonary lipid phosphate
RT phosphohydrolases.";
RL Am. J. Physiol. 281:L1484-L1493 (2001).
RN [3]
RP CHARACTERIZATION.
RX PubMed=8663293; DOI=10.1074/jbc.271.28.16506;
RA Waggoner D.W., Gomez-Munoz A., Dewald J., Brindley D.N.;
RT "Phosphatidate phosphohydrolase catalyzes the hydrolysis of ceramide
RT 1-phosphate, lysophosphatidate, and sphingosine 1-phosphate.";
RL J. Biol. Chem. 271:16506-16509 (1996).
CC -1- FUNCTION: Broad-specificity phosphohydrolase that dephosphorylates
CC exogenous bioactive glycerolipids and sphingolipids. Catalyzes the
CC conversion of phosphatidic acid (PA) to diacylglycerol (DG). In
CC addition it hydrolyzes lysophosphatidic acid (LPA), diacyl
CC glycerol pyrophosphate (DGPP), ceramide-1-phosphate (C-1-P) and
CC sphingosine-1-phosphate (S-1-P). The relative catalytic efficiency
CC is LPA > PA > C-1-P > S-1-P.
CC -1- CATALYTIC ACTIVITY: A 3-sn-phosphatidate + H(2)O = a 1,2-diacyl-
CC sn-glycerol + phosphate.
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Plasma membrane
CC (By similarity).
CC -1- ALTERNATIVE PRODUCTS:

```

```

CC Event=Alternative splicing; Named isoforms=2;
CC Comment=Additional isoforms seem to exist;
CC Name=1;
CC IsoId=O08564-1; Sequence=Displayed;
CC Name=2; Synonyms=Lppl;
CC IsoId=O08564-2; Sequence=VSP 009653;
CC -1- SIMILARITY: Belongs to the PA_phosphatase related phosphoesterase
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U90556; AAB50246.1; -.
CC EMBL; AF503609; AAM28631.1; -.
CC GO; GO:0005624; C:membrane fraction; ISS.
CC GO; GO:0008195; P:phosphatidate phosphatase activity; ISS.
CC GO; GO:0030521; P:androgen receptor signaling pathway; ISS.
CC GO; GO:0008285; P:negative regulation of cell proliferation; ISS.
CC GO; GO:0046839; P:phospholipid dephosphorylation; ISS.
CC GO; GO:0007205; P:protein kinase C activation; ISS.
CC GO; GO:0019216; P:regulation of lipid metabolism; ISS.
CC InterPro; IPR008934; AcPase_VanPerase.
CC Pfam; PF01569; PAP2; 1.
CC SMART; SM00014; acidPPC; 1.
KW Alternative splicing; Glycoprotein; Hydrolase; Transmembrane.
FT DOMAIN 1 6
FT TRANSMEM 7 27
FT DOMAIN 28 53
FT DOMAIN 54 74
FT TRANSMEM 75 88
FT DOMAIN 89 109
FT TRANSMEM 110 164
FT DOMAIN 165 185
FT TRANSMEM 186 194
FT DOMAIN 195 215
FT TRANSMEM 216 229
FT TRANSMEM 230 250
FT DOMAIN 251 282
FT CARBOHYD 142 142
FT VARSPIC 21 70
FT
FT
FT
SQ SEQUENCE 282 AA; 31996 MW; A4ED3DEB33FD7943 CRC64;
Query Match 60.6%; Score 66; DB 1; Length 282;
Best Local Similarity 50.0%; Pred. No. 0.015;
Matches 10; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 EGYIYNYRCRGDDSKVQEAR 20
DB 145 DGYIENYVCKGNAQKVEGR 164

RESULT 9
QGP766
ID Q6P766 PRELIMINARY; PRT; 282 AA.
AC Q6P766;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]

```

RP SEQUENCE FROM N.A.
 RC TISSUE=Prostate;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Bhat N.K.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S.S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny N.K., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [12]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Prostate;
 RA Strausberg R.;
 RL Submitted (NOV-2003) to the EMBL/GenBank/DBSJ databases.
 DR EMBL; BC061815; AH61815.1; -;
 DR InterPro; IPR008934; AcPase VanParase.
 DR InterPro; IPR00326; Pesterase_PA_PTP.
 DR Pfam; PF01569; PAP2; 1.
 DR SMART; SM00014; acidPPc; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 282 AA; 32000 MW; 10BFDD1C5F3FCB8E9 CRC64;
 Query Match 60.6%; Score 66; DB 2; Length 282;
 Best Local Similarity 50.0%; Pred. No. 0.015;
 Matches 10; Conservative 7; Mismatches 3; Indels 0; Gaps 0;
 QY 1 EGYIQNYRCRGDSKQVQEAR 20
 DB 145 DGYIENFVCCQNEQKVRREG 164
 RESULT 10
 LPPI_MOUSE
 ID LPPI_MOUSE STANDARD; PRT; 283 AA.
 AC Q61469; Q61690; Q8BPB8;
 DT 29-MAR-2004 (Rel. 43, Created)
 DT 29-MAR-2004 (Rel. 43, Last sequence update)
 DE 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Lipid phosphate phosphohydrolase 1 (EC 3.1.3.4) (Phosphatidic acid
 DE phosphatase 2a) (Phosphatidate phosphohydrolase type 2a) (PAP2a) (PAP-
 DE 2a) (PAP2-alpha) (Hydrogen peroxide-inducible protein 53) (Hic53) (35-
 DE kDa PAP) (MPAP).
 GN Name=Pap2a; Synonym=Hpic53, LPP1;
 OS Mus musculus (Mouse)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND INDUCTION.
 RC TISSUE=Calvaria;
 RX MEDLINE=96032549; PubMed=7556647; DOI=10.1016/0014-5793(95)00957-B;
 RA Egawa K., Yoshiwara M., Shibamura M., Nose K.;
 RT "Isolation of a novel ras-recision gene that is induced by hydrogen
 RT peroxide from a mouse osteoblastic cell line, MC3T3-E1.";
 RL FEBS Lett. 372:74-77 (1995).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Kidney;
 RX MEDLINE=96324980; PubMed=8702556; DOI=10.1074/jbc.271.31.18931;

RA Kai M., Wada I., Imai S.-I., Sakane F., Kanoh H.;
 RT "Identification and cDNA cloning of 35-kDa phosphatidic acid
 RT phosphatase (type 2) bound to plasma membranes.";
 RL J. Biol. Chem. 271:18931-18938 (1996).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC STRAIN=FVB/N;
 RA Yokoyama K., Tigyi G.;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBSJ databases.
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC STRAIN=C57BL/6J; TISSUE=Ovary, and Uterus;
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaaki I., Oeato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sautelina A., Schneider C., Sempile C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Wilmong L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573 (2002).
 RN [5]
 RP CHARACTERIZATION.
 RC TISSUE=Liver;
 RX PubMed=10359651; DOI=10.1042/0264-6021.3400677;
 RA Jasinska R., Zhang Q.-X., Pilquil C., Singh I., Xu J., Dewald J.,
 RA Dillon D.A., Berthiaume L.G., Carman G.W., Waggoner D.W.,
 RA Brindley D.N.;
 RT "Lipid phosphate phosphohydrolase-1 degrades exogenous glycerolipid
 RT and sphingolipid phosphate esters.";
 RL Biochem. J. 340:677-686 (1999).
 RN [6]
 RP SUBUNIT.
 RX PubMed=14725715; DOI=10.1186/1471-2091-5-2;
 RA Burnett C., Makridou P., Hewlett L., Howard K.;
 RT "Lipid phosphate phosphatases dimerise, but this interaction is not
 RT required for in vivo activity.";
 RL BMC Biochem. 5:2-2 (2004).
 RN [7]
 RP OVEREXPRESSION.
 RX PubMed=14687668; DOI=10.1016/j.cellsig.2003.08.012;
 RA Yue J., Yokoyama K., Balazs L., Baker D.L., Smalley D., Pilquil C.,
 RA Brindley D.N., Tigyi G.;
 RT "Mice with transgenic overexpression of lipid phosphate phosphatase-1
 RT display multiple organotypic deficits without alteration in
 RT circulating lysophosphatidate level.";
 RL Cell. Signal. 16:385-399 (2004).
 CC -!- FUNCTION: Broad-specificity phosphohydrolase that dephosphorylates
 CC exogenous bioactive glycerolipids and sphingolipids. Catalyzes the
 CC conversion of phosphatidic acid (PA) to diacylglycerol (DG). In
 CC addition it hydrolyzes lysophosphatidic acid (lPA), diacyl
 CC glycerol pyrophosphate (DGPP), ceramide-1-phosphate (C-1-P) and

CC sphingosine-1-phosphate (S-1-P). The relative catalytic efficiency
 CC is LPA > PA > C-1-P > S-1-P.
 CC -!- CATALYTIC ACTIVITY: A 3-sn-phosphatidate + H(2)O = a 1,2-diacyl-
 CC sn-glycerol + phosphate.
 CC -!- SUBUNIT: Homodimer. This complex seems not to be involved in
 CC substrate recognition, it may confer only structural or functional
 CC stability.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Found
 CC predominantly in plasma membrane.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q61469-1; Sequence=displayed;
 CC Name=2;
 CC IsoId=Q61469-2; Sequence=VSP_009652;
 CC -!- TISSUE SPECIFICITY: Highly expressed in kidney and lung. Almost
 CC undetectable in brain, heart, bone, muscle or spleen.
 CC -!- INDUCTION: Moderately, by hydrogen peroxide, calcium ionophore and
 CC dexamethasone.
 CC -!- PTM: N-glycosylated. Contains high-mannose oligosaccharide.
 CC -!- MISCELLANEOUS: Overexpression elicited a number of phenotypic
 CC alteration without affecting several aspects of LPA signaling.
 CC Phenotypic abnormalities affect primarily three organs: the liver,
 CC the skin, and the reproductive organs. There is a reduction on
 CC body size, birth weight, abnormalities in fur growth, and a
 CC severely impaired spermatogenesis.
 CC -!- SIMILARITY: Belongs to the PA-phosphatase related phosphoesterase
 CC family.
 CC -!- CAUTION: Ref.1 sequence differs from that shown due to numerous
 CC framehifts.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; LA3371; AA085353.1; AUT_SEQ.
 CC EMBL; D84376; BA012335.1; -.
 CC EMBL; AY247795; AAP04434.1; -.
 CC EMBL; AY247796; AAP04435.1; -.
 CC EMBL; AK07275; BAC36724.1; -.
 CC F01; S66668; S66668.
 CC MGD; MGI:104646; Hpic53.
 CC MGD; MGI:108412; Ppap2a.
 CC GO; GO:0005624; C:membrane fraction; ISS.
 CC GO; GO:0008195; P:phosphatidate phosphatase activity; ISS.
 CC GO; GO:0030521; P:androgen receptor signaling pathway; ISS.
 CC GO; GO:0008285; P:negative regulation of cell proliferation; ISS.
 CC GO; GO:0046839; P:phospholipid dephosphorylation; ISS.
 CC GO; GO:0006470; P:protein amino acid dephosphorylation; IDA.
 CC GO; GO:0007205; P:protein kinase C activation; ISS.
 CC GO; GO:0019216; P:regulation of lipid metabolism; ISS.
 CC InterPro; IPR008934; Acpase_vanPerase.
 CC InterPro; IPR000326; Pesterase_PA_FTP.
 CC Pfam; PF01569; PAP2; 1.
 CC SMART; SM00014; acidppc; 1.
 CC Alternative splicing: Glycoprotein; Hydrolase; Transmembrane.
 CC DOMAIN 1 6 Cytoplasmic (Potential).
 CC TRANSMEM 7 27 Potential.
 CC DOMAIN 28 53 Extracellular (Potential).
 CC TRANSMEM 54 74 Potential.
 CC DOMAIN 75 88 Cytoplasmic (Potential).
 CC TRANSMEM 89 109 Potential.
 CC DOMAIN 110 164 Extracellular (Potential).
 CC TRANSMEM 165 185 Potential.
 CC DOMAIN 186 199 Cytoplasmic (Potential).
 CC TRANSMEM 200 220 Potential.
 CC DOMAIN 221 229 Extracellular (Potential).
 CC TRANSMEM 230 250 Potential.
 CC DOMAIN 251 283 Cytoplasmic (Potential).

FT CARBOHYD 142 142 N-linked (GlcNAc...) (Potential).
 FT VARSPLIC 21 70
 FT IVPFCIIIV -> AMPMTILKLGKVPFQGFCTDINSVKY
 FT PYHDSITPSRIAILLGLPLIFS (in isoform 2).
 FT /FTID=VSP_009652.
 SQ SEQUENCE 283 AA; 31891 MW; 669690568549CC6 CRC64;
 Query Match 59.6%; Score 65; DB 1; Length 283;
 Best Local Similarity 50.0%; Pred. No. 0.022;
 Matches 10; Conservative 7; Mismatches 3; Indels 0; Gaps 0;
 QY 1 EGVIONYRCRGDDSKVQEAR 20
 DB 145 DGIYSDYICQGNKXKGR 164
 RESULT 11
 LPPI_CAVPO STANDARD; PRT; 285 AA.
 ID LPPI_CAVPO
 AC O88956; O88957;
 DT 29-MAR-2004 (Rel. 43, Created)
 DT 29-MAR-2004 (Rel. 43, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Lipid phosphate phosphohydrolase 1 (EC 3.1.3.4) (Phosphatidic acid
 DE phosphatase 2a) (Phosphatidate phosphohydrolase type 2a) (PAP2a) (PAP-
 DE 2a) (PAP2-alpha).
 GN Name=PPAP2A; Synonyms=LPP1, PAP2A;
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
 OX NCBI_TaxID=10141;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE=Airway smooth muscle;
 RX MEDLINE=99334234; PubMed=10405762; DOI=10.1016/S0898-6568(99)00028-5;
 RA Tate R.J., Tolan D., Pyne S.;
 RT "Molecular cloning of magnesium-independent type 2 phosphatidic acid
 RT phosphatases from airway smooth muscle.";
 RL Cell. Signal. 11:515-522(1999).
 CC -!- FUNCTION: Broad-specificity phosphohydrolase that dephosphorylates
 CC exogenous bioactive glycerolipids and sphingolipids. Catalyzes the
 CC conversion of phosphatidic acid (PA) to diacylglycerol (DG). In
 CC addition it hydrolyzes lysophosphatidic acid (LPA), ceramide-1-
 CC phosphate (C-1-P) and sphingosine-1-phosphate (S-1-P) (By
 CC similarity).
 CC -!- CATALYTIC ACTIVITY: A 3-sn-phosphatidate + H(2)O = a 1,2-diacyl-
 CC sn-glycerol + phosphate.
 CC -!- SUBUNIT: Homodimer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Plasma membrane
 CC (By similarity).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1; Synonyms=PAP2a1;
 CC IsoId=O88956-1; Sequence=Displayed;
 CC Name=2; Synonyms=PAP2a2;
 CC IsoId=O88956-2; Sequence=VSP_009650;
 CC -!- SIMILARITY: Belongs to the PA-phosphatase related phosphoesterase
 CC family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF088283; AAC63333.1; -.
 CC EMBL; AF088284; AAC63334.1; -.
 CC GO; GO:0005624; C:membrane fraction; ISS.
 CC GO; GO:0008195; P:phosphatidate phosphatase activity; ISS.
 CC GO; GO:0030521; P:androgen receptor signaling pathway; ISS.
 CC GO; GO:0008285; P:negative regulation of cell proliferation; ISS.

DR GO: GO:0046839; P:phospholipid dephosphorylation; ISS.
DR GO: GO:0007205; P:protein kinase C activation; ISS.
DR GO: GO:0019216; P:regulation of lipid metabolism; ISS.
DR InterPro: IPR008934; ACPase_VanPerase.
DR InterPro: IPR000326; Pesterase_PA_PTP.
DR Pfam: PF01569; PAP2; 1.
DR SMART: SM00014; acidPPC; 1.
KW Alternative splicing; Glycoprotein; Hydrolase; Transmembrane.
FT DOMAIN 1 6 Cytoplasmic (Potential).
FT TRANSMEM 7 27 Potential.
FT DOMAIN 28 53 Luminal (Potential).
FT TRANSMEM 54 74 Potential.
FT DOMAIN 75 94 Cytoplasmic (Potential).
FT TRANSMEM 95 115 Potential.
FT DOMAIN 116 165 Luminal (Potential).
FT TRANSMEM 166 186 Potential.
FT DOMAIN 187 199 Cytoplasmic (Potential).
FT TRANSMEM 200 220 Potential.
FT DOMAIN 221 229 Luminal (Potential).
FT TRANSMEM 230 250 Potential.
FT DOMAIN 251 285 N-linked (GlcNAc-) (Potential).
FT CARBOHYD 142 142 GLPFAILTSRHTPFGIFCNDESICYKYKEDTIPYALLGG
FT VARSPIC 21 70 PYHDSIV -> SMPVAVINIGQYPPFGFGFCNDNSIQY
FT CONFLICT 5 5 A -> T (in Ref. 1; AACG33334).
FT CONFLICT 218 218 I -> V (in Ref. 1; AACG33334).
FT SEQUENCE 285 AA; 32133 MW; E6F48E188DE6CF5 CRG64;
Query Match 56.9%; Score 62; DB 1; Length 285;
Best Local Similarity 55.0%; Pred. No. 0.069;
Matches 11; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
QY 1 EGYIQYRCRGDSKQVEAR 20
DB 145 DGYIEYVCRGNAEKVKEGR 164
RESULT 12
LP1_HUMAN
ID LP1_HUMAN STANDARD; PRT; 284 AA.
AC O1494; O60457; O60463;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Lipid phosphate phosphohydrolase 1 (EC 3.1.3.4) (Phosphatidic acid
DE phosphatase 2a) (Phosphatidate phosphohydrolase type 2a) (PAP2a) (PAP-
DE 2a) (PAP2-alpha).
GN Name=PPAP2A; Synonym=LPP1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A. (ISOFORM 1), AND CHARACTERIZATION.
RX MEDLINE=97450990; PubMed=9305923; DOI=10.1074/jbc.272.39.24572;
RA Kai M., Wada I., Imai S.-I., Sakane F., Kanoh H.;
RT "Cloning and characterization of two human isozymes of Mg2+-
RT independent phosphatidic acid phosphatase.";
RL J. Biol. Chem. 272:24572-24578 (1997).
[2]
RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Lung;
RX MEDLINE=98230079; PubMed=9570154;
RA Leung D.W., Tompkins C.K., White T.;
RT "Molecular cloning of two alternatively spliced forms of human
RT phosphatidic acid phosphatase cDNAs that are differentially expressed
RT in normal and tumor cells.";
RL DNA Cell Biol. 17:377-385 (1998).
[3]
RN SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Prostate;

RX MEDLINE=98136170; PubMed=9468526; DOI=10.1074/jbc.273.8.4660;
RA Ulrix W.E.J., Swinnen J., Heyns W., Verhoeven G.;
RT "Identification of the phosphatidic acid phosphatase type 2a isozyme
RT as an androgen-regulated gene in the human prostatic Adenocarcinoma
RT cell line LNCaP.";
RL J. Biol. Chem. 273:4660-4665 (1998).
[4]
RN SEQUENCE FROM N.A. (ISOFORM 1), AND CHARACTERIZATION.
RX MEDLINE=98371049; PubMed=9705349; DOI=10.1074/jbc.273.34.22059;
RA Roberts R., Sciorra V.A., Morris A.J.;
RT "Human type 2 phosphatidic acid phosphohydrolases. Substrate
RT specificity of the type 2a, 2b, and 2c enzymes and cell surface
RT activity of the 2a isoform.";
RL J. Biol. Chem. 273:22059-22067 (1998).
[5]
RN SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Boak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[6]
RN FUNCTION
RX PubMed=12909631; DOI=10.1074/jbc.M306709200;
RA Smyth S.S., Sciorra V.A., Sigal Y.J., Pamulkar Z., Wang Z., Xu Y.,
RA Prestwich G.D., Morris A.J.;
RT "Lipid phosphate phosphatases regulate lysophosphatidic acid
RT production and signaling in platelets: studies using chemical
RT inhibitors of lipid phosphate phosphatase activity.";
RL J. Biol. Chem. 278:43214-43223 (2003).
CC -!- FUNCTION: Broad-specificity phosphohydrolase that dephosphorylates
CC exogenous bioactive glycerolipids and sphingolipids. Catalyzes the
CC conversion of phosphatidic acid (PA) to diacylglycerol (DG).
CC Pivotal regulator of lysophosphatidic acid (LPA) signaling in the
CC cardiovascular system. Major enzyme responsible of
CC dephosphorylating LPA in platelets, which terminates signaling
CC actions of LPA. May control circulating, and possibly also
CC regulate localized, LPA levels resulting from platelet activation.
CC It has little activity towards ceramide-1-phosphate (C-1-P) and
CC sphingosine-1-phosphate (S-1-P). The relative catalytic efficiency
CC is LPA > PA > S-1-P > C-1-P. It's down-regulation may contribute
CC to the development of colon adenocarcinoma.
CC -!- CATALYTIC ACTIVITY: A 3-sn-phosphatidate + H(2)O = a 1,2-diacyl-
CC sn-glycerol + phosphate.
CC -!- ENZYME REGULATION: Inhibited by sphingosine, zinc ions and
CC propanolol. Not inhibited by N-ethylmaleimide treatment.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Plasma membrane.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Comment=Additional isoforms seem to exist;
CC Name=1; Synonyms=Alpha-1, hLPP1, PAP2-a1;
CC IsoId=O14494-1; Sequence=Displayed;
CC Name=2; Synonyms=Alpha-2, hLPP1-a, PAP2-a2;
CC IsoId=O14494-2; Sequence=VSP_009651;
CC -!- TISSUE SPECIFICITY: Ubiquitously expressed with highest expression

AC	G6P13;	
DT	05-JUL-2004 (TrEMBLrel. 27, Created)	
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)	
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)	
DE	MGC90748 protein.	
DE	Name:MGC80748;	
GN	Xenopus laevis (African clawed frog).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;	
OC	Xenopodinae; Xenopus.	
OX	NCBI_TaxID=8355;	
OX	[1]	
RN	SEQUENCE FROM N.A.	
RP	TISSUE=Spleen;	
RC	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.243603899;	
RC	Klausner R.L., Colling F.S., Wagner L., Shennen C.M., Schuler G.D.,	
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,	
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,	
RA	Diatchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong L.,	
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,	
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,	
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,	
RA	Rock S.A., McGraw P.J., McKernan K.J., Malek J.A., Gunarathne P.H.,	
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,	
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,	
RA	Fahey J., Helkon E., Kettman M., Madan A., Rodrigues S., Sanchez A.,	
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,	
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,	
RA	Blakesley A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,	
RA	Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,	
RA	Jones S.J., Marra M.A.;	
RT	"Generation and initial analysis of more than 15,000 full-length human	
RT	and mouse cDNA sequences."	
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).	
RL	[2]	
RN	SEQUENCE FROM N.A.	
RP	TISSUE=Spleen;	
RC	MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;	
RC	Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,	
RA	Richardson P.;	
RA	"Genetic and genomic tools for Xenopus research: The NIH Xenopus	
RT	initiative."	
RT	Dev. Dyn. 225:384-391(2002).	
RT	[3]	
RN	SEQUENCE FROM N.A.	
RP	TISSUE=Spleen;	
RC	Klein S., Gerhard D.S.;	
RL	Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.	
RL	EMBL; BC073336; AAH73336.1; "	
DR	InterPro; IPR008934; AcPase Vanperase.	
DR	InterPro; IPR000326; Pesterase_PA_PTP.	
DR	Pfam; PF01569; PAP2; 1.	
DR	SMART; SM00014; acd1PPC; 1.	
SQ	SEQUENCE 284 AA; 31834 MW; 30D0C7391BB424CB CRC64;	
	Query Match 48.6%; Score 53; DB 2; Length 284;	
	Best Local Similarity 47.4%; Pred.No.2.2;	
	Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps	
QY	2 GYIQNYRCRGDSDSKVQEAR 20	
DB	147 GYIETFCVCEGDPKTSSEGR 165	
	: :	
RESULT 14		
ID	_Y787 METJA STANDARD; PRT; 504 AA.	
AC	Q58197;	
DT	29-MAR-2004 (Rel. 43, Created)	
DT	29-MAR-2004 (Rel. 43, Last sequence update)	
DT	05-JUL-2004 (Rel. 44, Last annotation update)	
DE	Hypothetical protein MJ0787.	

```
OrderedLocusNames=MJ0787;
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii.";
RL Science 273:1058-1073(1996).
CC -!- SIMILARITY: Strong, to M.thermoautotrophicum MTH1137.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U67523; AAB98783.1; -.
DR PIR; C64398; C64398.
DR TIGR; MJ0787; -.
DR InterPro; IPR011060; Ribp_bind_barrel.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 504 AA; 56128 MW; 1F6C18C2C655SEC0 CRC64;
Query Match 47.7%; Score 52; DB 1; Length 504;
Best Local Similarity 57.9%; Pred. No. 6.2;
Matches 11; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 EGYIQNYRCRGDSDSKVQEA 19
| | | | | | | | | | | | | | | | | | | |
Db - 457 EAGINAYPCNGDDKKVLEA 475
-----
RESULT 15
Q95ZL6 PRELIMINARY; PRT; 244 AA.
AC Q95ZL6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein M01E11.4.
GN Name=M01E11.4; ORFNames=M01E11.4;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG WormBase Consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Pauley A., Gattung S.;
RT "The sequence of C. elegans cosmid M01E11.";
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RG WormBase Consortium;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; U80450; AAK77614.1; -.
```

```

RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RG WormBase Consortium;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; U80450; AAK77613.1; -.
DR WormBase; WBGene00019713; M01E11.4.
DR WormPep; M01E11.4a; CE28610.
KW Hypothetical protein.
SQ SEQUENCE 244 AA; 27837 MW; 9AA335B0442B84C4 CRC64;
Query Match 46.8%; Score 51; DB 2; Length 244;
Best Local Similarity 50.0%; Pred. No. 4.1;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
QY 1 EGYIQNYRCRGDSDSKVQEA 18
| | | | | | | | | | | | | | | | | | | |
Db 34 EGFVTRGGCRGLDSQIQE 51
-----
RESULT 16
Q95ZL7 PRELIMINARY; PRT; 252 AA.
AC Q95ZL7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein M01E11.4.
GN Name=M01E11.4; ORFNames=M01E11.4;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG WormBase Consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Pauley A., Gattung S.;
RT "The sequence of C. elegans cosmid M01E11.";
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RG WormBase Consortium;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; U80450; AAK77614.1; -.
```

```

DR WormBase: WBGene00019713; M01E11.4.
DR WormPep: M01E11.4b; CE28611.
KW Hypothetical protein.
SQ SEQUENCE 252 AA; 28738 MW; FE8556D34587F3E8 CRC64;

Query Match 46.8%; Score 51; DB 2; Length 252;
Best Local Similarity 50.0%; Pred. No. 4.2;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 EGYQNYRCRGDDSKVQE 18
||:| | | | | | | |
Db 34 EGVTRGCGRLDSQIQE 51

RESULT 17
Q7RTC3 PRELIMINARY; PRT; 2736 AA.
AC Q7RTC3;
DT 01-MAR-2004 (T-EMBLrel. 26, Created)
DT 01-MAR-2004 (T-EMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN Names=PY00071;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=73239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=17XNL;
RX PubMed=12368865; DOI=10.1038/nature01099;
RA Carlton J.M., Anguilo S.V., Suh B.B., Kooij T.W., Perteau M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA Shallow S.J., van Aken S.E., Riedmuller S.B., Feldlyum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoabi A., Cummings L.M.,
RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria
RT parasite Plasmodium yoelii yoelii."
RL Nature 419:512-519(2002).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABL0100019; EAA17884.1; -.
DR InterPro: IPR000357; HEAT.
DR Pfam: PF02985; HEAT; 1.
KW Hypothetical protein.
SQ SEQUENCE 2736 AA; 323770 MW; C219F762539AA00 CRC64;

Query Match 46.8%; Score 51; DB 2; Length 2736;
Best Local Similarity 50.0%; Pred. No. 58;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 YIQNYRCRGDDSKVQE 18
||:| | | | | | | |
Db 2533 YIYVACSTDQKID 2548

RESULT 18
LPP2_MOUSE
ID LPP2_MOUSE STANDARD; PRT; 276 AA.
AC Q9DAX2; Q9UW44;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 05-JUL-2004 (Rel. 43, Last annotation update)
DE Lipid phosphate phosphohydrolase 2 (EC 3.1.3.4) (Phosphatidic acid
DE phosphatase 2c) (PAP2-gamma) (PAP2-G).
DE 2c) (PAP2-gamma) (PAP2-G).
GN Names=Ppap2c; Synonyms=Lpp2;
OS Mus musculus (Mouse).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC STRAIN=C57BL/6J;
RC MEDLINE=20130123; PubMed=10662554; DOI=10.1006/geno.1999.6055;
RA Zhang N., Copeland N.G., Gilbert D.J., Jenkins N.A., Gridley T.;
RA "Cloning, expression, and chromosomal localization of a mouse gene
RT homologous to the germ cell migration regulator wunen and to type 2
RT phosphatidic acid phosphatases.";
RL Genomics 63:142-144(2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=C57BL/6J; TISSUE=Embryo, and Placenta;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Kanagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Mikki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Perteau G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sadelain A., Schneider C., Sempile C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Breast tumor;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.D., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smalish D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Catalyzes the conversion of phosphatidic acid (PA) to
CC diacylglycerol (DG). In addition it hydrolyzes lysophosphatidic
CC acid (LPA), ceramide-1-phosphate (C-1-P) and sphingosine-1-
CC phosphate (S-1-P) (By similarity).

```

CC -|- CATALYTIC ACTIVITY: A 3-sn-phosphatidate + H(2)O = a 1,2-diacyl-
 CC sn-glycerol + phosphate.
 CC -|- SUBUNIT: Homodimer (By similarity).
 CC -|- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -|- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q9DAX2-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q9DAX2-2; Sequence=VSP_009654, VSP_009655;
 CC -|- TISSUE SPECIFICITY: Expressed at high levels in lung, liver and
 CC kidney; at low levels in heart and brain, and was not detected in
 CC skeletal muscle.
 CC -|- SIMILARITY: Belongs to the PA-phosphatase related phosphoesterase
 CC family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF123611; AAD24061.1; -;
 CC EMBL; AK005452; BAB24045.1; -;
 CC EMBL; AK049581; BAC33824.1; -;
 CC EMBL; BC010332; AAH10332.1; -;
 CC MGD; MGI:1354945; Ppap2c.
 CC InterPro; IPR008934; AcPase_VanPerase.
 CC InterPro; IPR000326; Pesterase_PA_PTP.
 CC Pfam; PF01569; PAP2; 1.
 CC SMART; SM00014; acidppc; 1.
 CC Alternative splicing; Glycoprotein; Hydrolase; Transmembrane.
 CC TRANSMEM 5 25 Potential.
 CC TRANSMEM 52 72 Potential.
 CC TRANSMEM 88 108 Potential.
 CC TRANSMEM 162 182 Potential.
 CC TRANSMEM 190 210 Potential.
 CC TRANSMEM 219 239 Potential.
 CC CARBOHYD 139 139 N-linked (GlcNAc...) (Potential).
 CC VARSPLIC 240 250 RVSVDFKSRP -> SPTCLTHRLCF (in isoform
 CC 2).
 CC FT VARSPLIC 251 276 /FTID=VSP_009654.
 CC FT Missing (in isoform 2).
 CC FT /FTID=VSP_009655.
 CC SQ SEQUENCE 276 AA; 31193 MW; 978A83D468113D2 CRC64;
 CC
 CC Query Match 45.9%; Score 50; DB 1; Length 276;
 CC Best Local Similarity 52.6%; Pred. No. 6.9;
 CC Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
 CC
 CC QY 2 GYQNYRCRGDSDSKVQEAR 20
 CC ||:| ||| : |||
 CC Db 142 GYVQVEVCRGSPANVTAR 160
 CC
 CC RESULT 19
 CC LPP2_RAT STANDARD; PRT; 276 AA.
 CC ID Q8K593;
 CC AC Q8K593;
 CC DT 29-MAR-2004 (Rel. 43, Created)
 CC DT 29-MAR-2004 (Rel. 43, Last sequence update)
 CC DT 05-JUL-2004 (Rel. 44, Last annotation update)
 CC DE Lipid phosphate phosphohydrolase 2 (EC 3.1.3.4) (Phosphatidic acid
 CC phosphatase 2c) (PAP2-G).
 CC DE 2c) (PAP2-gamma) (PAP2-G).
 CC GN Name=Ppap2c; Synonyms=Lpp2;
 CC OS Rattus norvegicus (Rat).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CC NCBI_TaxID=10116;
 CC RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Lung;
 RA Zhao L., Nanjundan M., Possmayer F.;
 RT "Expression of lipid phosphate phosphohydrolase isoforms in rat lung
 RT development.";
 RL Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
 RN [2]
 RN TISSUE SPECIFICITY.
 RP STRAIN=Sprague-Dawley;
 RX MEDLINE=21559999; PubMed=11704545;
 RA Nanjundan M., Possmayer F.;
 RT "Molecular cloning and expression of pulmonary lipid phosphate
 RT phosphohydrolases.";
 RL Am. J. Physiol. 281:L1484-L1493 (2001).
 CC -|- FUNCTION: Catalyzes the conversion of phosphatidic acid (PA) to
 CC diacylglycerol (DG). In addition it hydrolyzes lysophosphatidic
 CC acid (LPA), ceramide-1-phosphate (C-1-P) and sphingosine-1-
 CC phosphate (S-1-P) (By similarity).
 CC -|- CATALYTIC ACTIVITY: A 3-sn-phosphatide + H(2)O = a 1,2-diacyl-
 CC sn-glycerol + phosphate.
 CC -|- SUBUNIT: Homodimer (By similarity).
 CC -|- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -|- TISSUE SPECIFICITY: Expressed in the brain.
 CC -|- SIMILARITY: Belongs to the PA-phosphatase related phosphoesterase
 CC family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF503611; AAM28632.1; -;
 CC InterPro; IPR008934; AcPase_VanPerase.
 CC InterPro; IPR000326; Pesterase_PA_PTP.
 CC Pfam; PF01569; PAP2; 1.
 CC SMART; SM00014; acidppc; 1.
 CC Glycoprotein; Hydrolase; Transmembrane.
 CC TRANSMEM 5 25 Potential.
 CC TRANSMEM 52 72 Potential.
 CC TRANSMEM 88 108 Potential.
 CC TRANSMEM 162 182 Potential.
 CC TRANSMEM 190 210 Potential.
 CC TRANSMEM 226 246 Potential.
 CC CARBOHYD 139 139 N-linked (GlcNAc...) (Potential).
 CC CARBOHYD 155 155 N-linked (GlcNAc...) (Potential).
 CC SQ SEQUENCE 276 AA; 31101 MW; 41A0E1243458E24C CRC64;
 CC
 CC Query Match 45.9%; Score 50; DB 1; Length 276;
 CC Best Local Similarity 52.6%; Pred. No. 6.9;
 CC Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
 CC
 CC QY 2 GYQNYRCRGDSDSKVQEAR 20
 CC ||:| ||| : |||
 CC Db 142 GYVQVEVCRGSPANVTAR 160
 CC
 CC RESULT 20
 CC Q83ER2 PRELIMINARY; PRT; 130 AA.
 CC ID Q83ER2;
 CC AC Q83ER2;
 CC DT 01-JUN-2003 (TReMBLrel. 24, Created)
 CC DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
 CC DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 CC DE Ribosomal protein S8.
 CC GN Name=rpsH; OrderedLocusNames=CBU0252;
 CC OS Coccidia burnetii.
 CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
 CC OC Cocciliaceae; Coccidia.
 CC NCBI_TaxID=777;
 CC RN [1]

[illegible]

RC TISSUE=EMBRYO;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;


```

RA Robison K.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
CC family.
DR EMBL; U00024; AAA50930.1; -.
DR HSP; P08659; 1LCI.
DR GO; GO:0003824; P:catalytic activity; IEA.
DR GO; GO:0048037; F:cofactor binding; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR009081; ACP like.
DR InterPro; IPR00873; AMP-bind.
DR InterPro; IPR006163; Phosphateth_bind.
DR Pfam; PF00501; AMP-binding; 1.
DR Pfam; PF00550; PP-binding; 1.
DR PRINTS; PR00154; AMPBINDING.
DR PROSITE; PS00075; ACP_DOMAIN; 1.
SQ SEQUENCE 705 AA; 75169 MW; 9f9d4f500690bf5d CRC64;

Query Match 42.7%; Score 46.5; DB 2; Length 705;
Best Local Similarity 50.0%; Pred. No. 74;
Matches 8; Conservative 5; Mismatches 2; Indels 1; Gaps 1;

QY 1 EGYQNYRCRGDDSKV 16
   :|:: |||| ||::|
DB 378 DGWV-TYRCRADDTTEV 392

RESULT 27
P96283
ID P96283 PRELIMINARY; PRT; 705 AA.
AC P96283; Q7D6D9;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE PROBABLE FATTY-ACID-CoA LIGASE FAD22 (FATTY-ACID-CoA SYNTHETASE)
DE (FATTY-ACID-CoA SYNTHASE) (PC 6.2.1.-) (Substrate--CoA ligase).
GN Name-fad22; OrderedLocNames=WT3021, RV2948c;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,
RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jagsis K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares R.,
RA Sultston J.E., Taylor K., Whitehead S., Barrrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Unayama L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RA "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL J. Bacteriol. 184:5479-5490 (2002).
CC -!- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
CC family.
DR EMBL; BX842581; CAB06101.1; -.

```

```

DR EMBL; AB000516; AAK47347.1; -.
DR PIR; A70669; A70669.
DR HSP; P08659; 1LCI.
DR TIGR; MT3021; -.
DR TubercuList; RV2948c; -.
DR GO; GO:0003824; P:catalytic activity; IEA.
DR GO; GO:0048037; F:cofactor binding; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR009081; ACP like.
DR InterPro; IPR00873; AMP-bind.
DR InterPro; IPR006163; Phosphateth_bind.
DR Pfam; PF00501; AMP-binding; 1.
DR Pfam; PF00550; PP-binding; 1.
DR PRINTS; PR00154; AMPBINDING.
DR PROSITE; PS00075; ACP_DOMAIN; 1.
KW Ligase.
SQ SEQUENCE 705 AA; 75197 MW; 600f2d0eabfddf1dc CRC64;

Query Match 42.7%; Score 46.5; DB 2; Length 705;
Best Local Similarity 50.0%; Pred. No. 74;
Matches 8; Conservative 5; Mismatches 2; Indels 1; Gaps 1;

QY 1 EGYQNYRCRGDDSKV 16
   :|:: |||| ||::|
DB 378 DGWV-TYRCRADDTTEV 392

RESULT 28
Q7TXK7
ID Q7TXK7 PRELIMINARY; PRT; 705 AA.
AC Q7TXK7;
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE PROBABLE FATTY-ACID-CoA LIGASE FAD22 (FATTY-ACID-CoA SYNTHETASE)
DE (FATTY-ACID-CoA SYNTHASE) (EC 6.2.1.-).
GN Name-fad22; OrderedLocNames=MB2972c;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Mensepe C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis."
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).
DR EMBL; BX248344; CAD96659.1; -.
DR GO; GO:0048037; F:cofactor binding; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR009081; ACP like.
DR InterPro; IPR00873; AMP-bind.
DR InterPro; IPR006163; Phosphateth_bind.
DR Pfam; PF00501; AMP-binding; 1.
DR Pfam; PF00550; PP-binding; 1.
DR PROSITE; PS00075; ACP_DOMAIN; 1.
KW Complete proteome; Ligase.
SQ SEQUENCE 705 AA; 75197 MW; 600f2d0eabfddf1dc CRC64;

Query Match 42.7%; Score 46.5; DB 2; Length 705;
Best Local Similarity 50.0%; Pred. No. 74;
Matches 8; Conservative 5; Mismatches 2; Indels 1; Gaps 1;

QY 1 EGYQNYRCRGDDSKV 16
   :|:: |||| ||::|
DB 378 DGWV-TYRCRADDTTEV 392

```

```

RESULT 29
Q6CZY4
ID Q6CZY4 PRELIMINARY; PRT; 130 AA.
AC Q6CZY4;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE 30S ribosomal subunit protein S8.
GN Name=rpsh; OrderedLocNames=ECA4017;
OS Erwinia carotovora (subsp. atroseptica) (Pectobacterium atrosepticum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=29471;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SCRI 1043 / ATCC BAA-672;
RX PubMed=15263089; DOI=10.1073/pnas.0402424101;
RA Bell K.S., Sebahia M., Pritchard L., Holden M.T.G., Hyman L.J.,
RA Holeva M.C., Thomson N.R., Bentley S.D., Churcher L.J.C., Mungall K.,
RA Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,
RA Fraser A., Hance Z., Hauser H., Jagels K., Moule S., Norbertczak H.,
RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
RA Salmon G.P.C., Birch P.R.J., Parkhill J., Toth I.K.;
RA "Genome sequence of the enterobacterial phytopathogen Erwinia
RT carotovora subsp. atroseptica and characterization of virulence
RT factors.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:11105-11110(2004).
CC -!- SIMILARITY: Belongs to the S8P family of ribosomal proteins.
DR EMBL; BX950851; -
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0005840; C:ribosome; IEA.
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR GO; GO:0006412; F:protein biosynthesis; IEA.
DR InterPro; IPR000630; Ribosomal S8.
DR Pfam; PF00410; Ribosomal S8; 1.
DR ProDom; PD001098; Ribosomal S8; 1.
DR PROSITE; PS00053; RIBOSOMAL_S8; 1.
KW Complete proteome; Ribonucleoprotein; Ribosomal protein.
SQ SEQUENCE 130 AA; 14211 MW; 5CBED0F565CD3A3C4 CRC64;

Query Match 42.2%; Score 46; DB 2; Length 130;
Best Local Similarity 44.4%; Pred. No. 14;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 EGYIYQYRCRGDSSKQVE 18
Db |||::: |||
43 EGYIEDFKIEGTPKPVLE 60

RESULT 30
Q7YU44
ID Q7YU44 PRELIMINARY; PRT; 340 AA.
AC Q7YU44;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE RE35739p.
GN ORFNames=CG11426;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champagne M., Chavez C., Dorsett V., Dreanek D., Farfan D., Friese E.,
RA George R., Gonzalez C., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;

```

```

RL Submitted (AUG-2003) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BT009997; AAQ22466.1; -
DR FlyBase; FBgn0037166; CG11426.
DR InterPro; IPR008934; AcPase VanParase.
DR InterPro; IPR00326; Pesterase_PA_PTP.
DR Pfam; PF01569; PAP2; 1.
DR SMART; SM00014; acidPPG; 1.
SQ SEQUENCE 340 AA; 37628 MW; 08FA16CDICE34636 CRC64;

Query Match 42.2%; Score 46; DB 2; Length 340;
Best Local Similarity 38.9%; Pred. No. 41;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 3 YIONVYRCRGDSSKQVEAR 20
Db |||::: |||
186 YNEDYDCAGEGFTVEDVR 203

RESULT 31
Q9VNT9
ID Q9VNT9 PRELIMINARY; PRT; 340 AA.
AC Q9VNT9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CG11426-PA.
GN ORFNames=CG11426;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yeandle M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brattier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cui Y., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foeller C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstock G.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Williams S.M., Woodage, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhou S., Zhu S.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
Science 287:2185-2195(2000).

```

```

RN  SEQUENCE FROM N.A.
RP  MEDLINE=22426065; PubMed=12537568;
RA  Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA  Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA  George R.A., Hoekins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA  Pacלב J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA  Svirskaas R., Tabor P.B., Wan K., Scapleton M., Sutton G.G., Venter C.,
RA  Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT  "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT  melanogaster euchromatic genome sequence.";
RL  Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN  [3]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=22426070; PubMed=12537573;
RA  Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskaas R.,
RA  Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA  Ashburner M., Celniker S.E.;
RT  "The transposable elements of the Drosophila melanogaster euchromatin:
RT  a genomics perspective.";
RL  Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN  [4]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=22426069; PubMed=12537572;
RA  Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA  Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA  Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA  Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA  Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA  Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA  Lewis S.E.;
RT  "Annotation of the Drosophila melanogaster euchromatic genome: a
RT  systematic review.";
RL  Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN  [5]
RP  SEQUENCE FROM N.A.
RG  FlyBase;
RL  Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN  [6]
RP  SEQUENCE FROM N.A.
RG  FlyBase;
RL  Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AS003597; AAF51831.1; -.
DR  FlyBase; FBgn0037166; CG11426.
DR  InterPro; IPR008934; AcPase_VanParase.
DR  InterPro; IPR000326; Pestertase_PA_PTP.
DR  Pfam; PF01569; PAP2; 1.
DR  SMART; SM00014; acidppc; 1.
SQ  SEQUENCE 340 AA; 37654 MW; 18E77AA8008F2323 CRC64;

Query Match 42.2%; Score 46; DB 2; Length 340;
Best Local Similarity 38.9%; Pred. NO. 41;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 3 YIQNYRCRGDDSKVQEAR 20
DB 186 YMENYDCAGEGFTVEDVR 203

RESULT 32
Q73RY0 PRELIMINARY; PRT; 776 AA.
AC Q73RY0
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=WAP4298c;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1770;
RN [1]

```

```

RP  SEQUENCE FROM N.A.
RC  STRAIN=K10;
RL  Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AE017242; AAS06848.1; -.
DR  GO; GO:0008152; P:metabolism; IEA.
DR  InterPro; IPR003673; CAIB_BAIF.
DR  Pfam; PF02515; CoA_transf_3; 1.
SQ  Complete proteome; Hypothetical protein.
KW SEQUENCE 776 AA; 82189 MW; 38A4528FCSB084BB CRC64;

Query Match 42.2%; Score 46; DB 2; Length 776;
Best Local Similarity 80.0%; Pred. NO. 18+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 YRCRGDDSKV 16
DB 615 YRCRGDDAW 624

RESULT 33
Q95Z20 PRELIMINARY; PRT; 1325 AA.
ID Q95Z20
AC Q95Z20
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein C09F9.3.
GN ORFNames=C09F9.3;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=Bristol N2;
RX  MEDLINE=99069613; PubMed=9851916;
RA  none;
RT  "Genome sequence of the nematode C.elegans: A platform for
RT  investigating biology.";
RL  Science 282:2012-2018(1998).
RN [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=Bristol N2;
RA  Smye R.;
RL  Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR  EMBL; Z81465; CAC42253.2; -.
DR  HSSP; O31465; 1MKT.
DR  WormBase; WBGene0007480; C09F9.3.
DR  WormPep; C09F9.3; C836977.
DR  GO; GO:0004359; F:glutaminase activity; IEA.
DR  GO; GO:0006541; P:glutamine metabolism; IEA.
DR  InterPro; IPR002110; ANK.
DR  InterPro; IPR007043; Glutaminase.
DR  Pfam; PF00023; Ank; 4.
DR  Pfam; PF04960; Glutaminase; 2.
DR  SMART; SM00248; ANK; 4.
DR  PROSITE; PS50088; ANK_REPEAT; 2.
DR  PROSITE; PS50297; ANK_REPEAT_REGION; 2.
KW ANK repeat; Hypothetical protein.
SQ  SEQUENCE 1325 AA; 150066 MW; D163D1D72D33BAC7 CRC64;

Query Match 42.2%; Score 46; DB 2; Length 1325;
Best Local Similarity 50.0%; Pred. NO. 1.8e+02;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 YIQNYRCRGDDSKVQES 18
DB 360 YMKYRCFPDESQVYE 375

RESULT 34
Q20683

```

```
ID Q20683 PRELIMINARY; PRT; 179 AA.
AC Q20683;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein F52H3.5.
GN ORFNames=F52H3.5;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
[2]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Gardner A.E.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z66512; CAA91325.1; -.
DR PIR; T22521; T22521.
DR WormBase; WEGene0000947; F52H3.5.
DR WormPep; F52H3.5; CE03401.
DR InterPro; IPR001440; TPR.
DR InterPro; IPR008941; TPR-like.
DR Pfam; PF00515; TPR 1; 1.
DR SMART; SM00028; TPR; 3.
DR PROSITE; PS50005; TPR; 1.
DR PROSITE; PS50293; TPR_REGION; 1.
DR Hypothetical protein; Repeat; TPR repeat.
KW Hypothetical protein; Repeat; TPR repeat.
SQ SEQUENCE 179 AA; 19648 MW; 6FF899837C65A18F CRC64;

Query Match 41.7%; Score 45.5; DB 2; Length 179;
Best Local Similarity 45.0%; Pred. No. 24;
Matches 9; Conservative 3; Mismatches 5; Indels 3; Gaps 1;

QY 1 EGYQYQ---YRCRGDSSKQV 17
Db 116 QAYVQRASIYRLRGDDDKAR 135

RESULT 35
Q6XNK7 PRELIMINARY; PRT; 1086 AA.
AC Q6XNK7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE NAD dependent glutamate dehydrogenase (EC 1.4.1.2).
GN Name=gdhB;
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=162425;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=G001;
RA Kawaguchi T., Unkles S., Kinghorn J.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY223544; AAP9491.1; -.
DR GO; GO:0004352; F:glutamate dehydrogenase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR InterPro; IPR006096; GLFV dehydrog_C.
DR Pfam; PF00208; GLFV_dehydrog; 1.
DR Oxidoreductase.
KW Oxidoreductase.
SQ SEQUENCE 1086 AA; 122429 MW; AF619A4384CC665E CRC64;
```

```
Query Match 41.7%; Score 45.5; DB 2; Length 1086;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 9; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

QY 1 EGYQYQCRGDDSKQVE 18
Db 947 QGFVQN-MCVGDDDSVPE 963

RESULT 36
Q7MYG5 PRELIMINARY; PRT; 130 AA.
AC Q7MYG5;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE 30S ribosomal protein S14.
GN Name=rpS14; OrderedLocusNames=plu4712;
OS Photorhabdus luminescens (subsp. laumondii).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Photorhabdus.
OX NCBI_TaxID=141679;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=TT01;
RX MEDLINE=22957627; PubMed=14528314;
RA Duchaud E., Rusniok C., Frangeul L., Buchrieser C., Givaudan A.,
RA Taourit S., Bocs S., Boursaux-Eude C., Chandler M., Charles J.-F.,
RA Dassa E., Derose R., Derzelle S., Freyssinet G., Gaudriault S.,
RA Medigue C., Lanois A., Powell K., Siguier P., Vincent R., Wingate V.,
RA Zouine M., Glaser P., Boemare N., Danchin A., Kunst F.;
RT "The genome sequence of the entomopathogenic bacterium Photorhabdus
RT luminescens.";
RL Nat. Biotechnol. 21:1307-1313(2003).
CC -!- SIMILARITY: Belongs to the S8P family of ribosomal proteins.
DR EMBL; BX571874; CAE17084.1; -.
DR HSSP; P24319; 1AN7.
DR PhotoList; plu4712; -.
DR GO; GO:0005840; C:ribosome; IEA.
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR InterPro; IPR000630; Ribosomal_S8.
DR Pfam; PF00410; Ribosomal_S8; 1.
DR ProDom; PD001098; Ribosomal_S8; 1.
DR PROSITE; PS00053; RIBOSOMAL_S8; 1.
DR Complete proteome; Ribonucleoprotein; Ribosomal protein.
KW Complete proteome; Ribonucleoprotein; Ribosomal protein.
SQ SEQUENCE 130 AA; 14215 MW; 37DEE66209DEC1C CRC64;

Query Match 41.3%; Score 45; DB 2; Length 130;
Best Local Similarity 58.3%; Pred. No. 21;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EGYQYQCRGD 12
Db 43 EGYIEDYKIEGD 54

RESULT 37
Q7RA96 PRELIMINARY; PRT; 300 AA.
AC Q7RA96;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative yir4 protein.
GN Name=PY06606;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=73239;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=17XNL;
RX PubMed=12368865; DOI=10.1038/nature01099;
```

Qy	1	EGYQNYRCRGDSDKV 16
Ds	37	DGYCDSYRCDNDFEKI 52
RESULT 39		
Q7RS76		
ID	Q7RS76	PRELIMINARY; PRT; 314 AA.
AC	Q7RS76;	
DT	01-MAR-2004	(TrEMBLrel. 26, Created)
DT	01-MAR-2004	(TrEMBLrel. 26, Last sequence update)
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)
DE	Putative yir4 protein.	
GN	Name=PY00488;	
OS	Plasmodium yoelii yoelii.	
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.	
OX	NCBI_TaxID=73239;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=17XNL;	
EX	PubMed=12368865; DOI=10.1038/nature01099;	
RA	Carlton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Pertea M.,	
RA	Silva J.C., Ermolaeva M.D., Allen J.E., Seielung J.D., Koo H.L.,	
RA	Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,	
RA	Shallom S.J., van Aken S.B., Riedmuller S.B., Feldblum T.V.,	
RA	Cho J.K., Quackenbush J., Sedegah M., Shoaihi A., Cummings L.M.,	
RA	Florens L., Yates F.R. III, Raine J.D., Sindén R.E., Harris M.A.,	
RA	Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,	
RA	van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,	
RA	Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,	
RA	Carucci D.J.;	
RT	"Genome sequence and comparative analysis of the model rodent malaria	
RL	parasite Plasmodium yoelii yoelii.";	
RL	Nature 419:512-519(2002).	
CC	-!- CAUTION: The sequence shown here is derived from an	
CC	EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is	
CC	preliminary data.	
DR	EMBL; AABL01000136; EAA16343.1; -.	
DR	InterPro; IPR006477; Yir_bir_cir.	
DR	Pfam; PF06032; Cir_Bir_Yir; I.	
DR	TIGRFAMs; TIGR01590; Yir-bir-cir_Pla; 1.	
SQ	SEQUENCE 314 AA; 36643 MW; F5E8D0BF0DCBIA1 CRC64;	
Query Match	41.3%; Score 45; DB 2; Length 314;	
Best Local Similarity	43.8%; Pred.No. 55;	
Matches	7; Conservative 3; Mismatches 6; Indels 0; Gaps	
Qy	1	EGYQNYRCRGDSDKV 16
Ds	39	DGYCDSYRCDNDFEKI 54
RESULT 40		
Q9AEN7		
ID	Q9AEN7	PRELIMINARY; PRT; 344 AA.
AC	Q9AEN7;	
DT	01-JUN-2001	(TrEMBLrel. 17, Created)
DT	01-JUN-2001	(TrEMBLrel. 17, Last sequence update)
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)
DE	Aspartate-B-semialdehyde dehydrogenase.	
GN	Name=asd;	
OS	Rhizobium sp. (strain NGR234).	
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;	
OC	Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.	
OX	NCBI_TaxID=394;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RA	Izallalen M., Levesque R.C., Perret X., Broughton W.J., Antoun H.;	
RL	Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.	
DR	EMBL; AY029475; AAC3122.1; -.	
DR	HSP; P44801; INWC.	
DR	GO; GO:0004073; F:aspartate-semialdehyde dehydrogenase activity; IEA:	

```

DR GO: GO:0006520; P:amino acid metabolism; IEA.
DR GO: GO:0009086; P:methionine biosynthesis; IEA.
DR GO: GO:0009088; P:threonine biosynthesis; IEA.
DR InterPro: IPR005986; Asp_ADH_USG1.
DR InterPro: IPR000534; Semiald_dehydrog.
DR Pfam: PF01118; Semialdehyde_dh; 1.
DR Pfam: PF02774; Semialdehyde_dhC; 1.
DR TIGRFAMs: TIGR01296; asd B; 1.
SQ SEQUENCE 344 AA; 37431 MW; EA1FF2BA0F742760 CRC64;

Query Match 41.3%; Score 45; DB 2; Length 344;
Best Local Similarity 38.1%; Pred. No. 61;
Matches 8; Conservative 5; Mismatches 4; Indels 4; Gaps 1;

QY 2 GYIQNYRCRGDD----SKVQE 18
Db 279 GYVTPYECAGEDATYISRIRE 299

RESULT 41
ID Q92KY2 PRELIMINARY; PRT; 344 AA.
AC Q92KY2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PUTATIVE ASPARTATE-SEMIALDEHYDE DEHYDROGENASE PROTEIN (EC
1.2.1.11).
GN ORFNames=Smc04410;
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430; DOI=10.1073/pnas.161294398;
RA Capela D., Barloy-Hubler J., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kias E., Lelaure V., Masuy D.,
RA Fohl T., Portetelle D., Puchler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.;
RA "Analysis of the chromosome sequence of the legume symbiont
RT - Sinorhizobium meliloti strain 1021.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882 (2001).
DR EMBL: AL591793; CAC47895.1; -.
DR HSSP: P44801; INWC.
DR GO: GO:0004073; F:aspartate-semialdehyde dehydrogenase activity; IEA.
DR GO: GO:0016491; F:oxidoreductase activity; IEA.
DR GO: GO:0006520; P:amino acid metabolism; IEA.
DR GO: GO:0009086; P:methionine biosynthesis; IEA.
DR GO: GO:0009088; P:threonine biosynthesis; IEA.
DR InterPro: IPR005986; Asp_ADH_USG1.
DR InterPro: IPR000534; Semiald_dehydrog.
DR Pfam: PF01118; Semialdehyde_dh; 1.
DR Pfam: PF02774; Semialdehyde_dhC; 1.
DR TIGRFAMs: TIGR01296; asd B; 1.
KW Complete proteome; Oxidoreductase.
SQ SEQUENCE 344 AA; 37573 MW; 0C8CF8F08B8D7A84 CRC64;

Query Match 41.3%; Score 45; DB 2; Length 344;
Best Local Similarity 38.1%; Pred. No. 61;
Matches 8; Conservative 5; Mismatches 4; Indels 4; Gaps 1;

QY 2 GYIQNYRCRGDD----SKVQE 18
Db 279 GYVTPYECAGEDATYISRIRE 299

RESULT 42
Q96I80 PRELIMINARY; PRT; 505 AA.
ID Q96I80
AC Q96I80;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PRSS12 protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]_
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC007761; AAH07761.1; -.
DR HSSP: O08380; 1BY2.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0005507; F:copper ion binding; IEA.
DR GO: GO:0004720; F:protein-lysine 6-oxidase activity; IEA.
DR GO: GO:0005044; F:scavenger receptor activity; IEA.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001695; Lysyl oxidase.
DR InterPro: IPR001190; Strc_receptor.
DR Pfam: PF00051; Kringle; 1.
DR Pfam: PF00530; SRCR; 3.
DR PRINTS: PR00018; KRINGLE.
DR PRINTS: PR00258; SPERACTRCPTR.
DR ProDom: PD000395; Kringle; 1.
DR ProDom: PD013887; Lysyl_oxidase; 1.
DR SMART: SM00130; KR; 1.
DR SMART: SM00202; SR; 3.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS00070; KRINGLE_2; 1.
DR PROSITE: PS00420; SRCR_1; 2.
DR PROSITE: PS00287; SRCR_2; 3.
KW Kringle.
SQ SEQUENCE 505 AA; 55658 MW; 7FF863A6246226BD CRC64;

Query Match 41.3%; Score 45; DB 2; Length 505;
Best Local Similarity 50.0%; Pred. No. 92;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 YIQNYRCGDDSKV 16
Db 233 YWSNVRCRGDEENI 246

RESULT 43
Q73606 PRELIMINARY; PRT; 607 AA.
ID P73606

```

```

AC P73606;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE S1r1855 protein.
GN OrderedLocusNames=slr1855;
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1149;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PCC6803;
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirose M., Sugita M., Sasaki S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
DR EMBL; D90908; BAA17650.1; -.
DR PIR; S77092; S77092.
DR InterPro; IPR008928; Glyco_trans_6hp.
KW Complete proteome.
SQ SEQUENCE 607 AA; 70003 MW; 873542F37B6AB0F3 CRC64;

Query Match 41.3%; Score 45; DB 2; Length 607;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 GYIQNYRCRGDPSKVQEA 19
DB 241 GPIQTYRCGTGDRRLHDA 258

RESULT 44
NETR HUMAN
ID NETR HUMAN STANDARD; PRT; 875 AA.
AC P56730; Q9UP16;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Neurotysin precursor (SC 3.4.21.-) (Motopsin) (Leydin).
GN Name=PRSS12;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98201705; PubMed=9540828; DOI=10.1016/S0167-4781(97)00205-4;
RA Proba K., Gschwend T.P., Sonderegger P.;
RT "Cloning and sequencing of the cDNA encoding human neurotysin.";
RL Biochim. Biophys. Acta 1396:143-147(1998).
RN [2]
RP SEQUENCE OF 615-875 FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=99203523; PubMed=101030556;
RA Poorafshar M., Hellman L.;
RT "Cloning and structural analysis of leydin, a novel human serine
RT protease expressed by the Leydig cells of the testis.";
RL Eur. J. Biochem. 261:244-250(1999).
RN [3]
RP DISEASE.
RX PubMed=12459588; DOI=10.1126/science.1076521;
RA Molinari F., Rio M., Meskenaitė V., Encha-Razavi F., Auge J., Bacq D.,
RA Briault S., Vekemans M., Munnich A., Attie-Bitach T., Sonderegger P.,
RA Collea L.;
RT "Truncating neurotysin mutation in autosomal recessive nonsyndromic
RT mental retardation.";
```

```

Science 298:1779-1781(2002).
-!- FUNCTION: Plays a role in neuronal plasticity and the proteolytic
CC action may subserve structural reorganizations associated with
CC learning and memory operations (By similarity).
-!- SUBCELLULAR LOCATION: Secreted.
CC
-!- TISSUE SPECIFICITY: Brain and Leydig cells of the testis.
CC
-!- DISEASE: Defects in PRSS12 are a cause of autosomal recessive
CC nonsyndromic mental retardation (MIM:249500). Mental retardation
CC is a mental disorder characterized by significantly subaverage
CC general intellectual functioning associated with impairments in
CC adaptive behavior and manifested during the developmental
CC period.
-!- SIMILARITY: Belongs to the peptidase S1 family.
CC
-!- SIMILARITY: Contains 1 kringle domain.
CC
-!- SIMILARITY: Contains 4 SRCR domains.
CC
This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
EMBL; AJ001531; CAA04816.1; -.
DR EMBL; AF077298; AAD25919.1; -.
DR HSP; P00760; IEZX.
DR MEROPS; S01.237; -.
DR Genew; HGNC:9477; PRSS12.
DR MIM; 606709; -.
DR MIM; 249500; -.
DR GO; GO:0008236; F:serine-type peptidase activity; TAS.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00530; SRCR; 4.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00258; SPERACTRCPTR.
DR ProDom; PD000395; Kringle; 1.
DR ProDom; PD013887; Lysyl_oxidase; 2.
DR SMART; SM00130; KR; 1.
DR SMART; SM00202; SR; 4.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00420; SRCR_1; 3.
DR PROSITE; PS0287; SRCR_2; 4.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Repeat; Serine protease; Signal.
FT SIGNAL 1 20
FT CHAIN 21 875
FT DOMAIN 23 92
FT DOMAIN 93 165
FT DOMAIN 170 271
FT DOMAIN 280 381
FT DOMAIN 387 487
FT DOMAIN 500 601
FT DOMAIN 619 875
FT DOMAIN 619 630
FT SITE 630 631
FT ACT_SITE 676
FT ACT_SITE 726
FT ACT_SITE 825
FT ACT_SITE 859
FT DISULFID 26 26
FT CARBOHYD 683 683
FT CARBOHYD 683 683
FT CONFLICT 663 663
```

Query Match 40.4%; Score 44; DB 2; Length 172;


```
Best Local Similarity 37.9%; Pred. No. 42;
Matches 11; Conservative 4; Mismatches 2; Indels 12; Gaps 2;

Qy 1 EGYI-----QNYRCRG-----DDSKVQ 17
Db 97 DGIAGIFVDSNYRCQIGLALDEAKKQ 125

RESULT 48
Q8SSM8 ID Q8SSM8 PRELIMINARY; PRT; 338 AA.
AC Q8SSM8
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2004 (TrEMBLrel. 21, Last sequence update)
DE FRUCTOSE-BISPHOSPHATE ALDOLASE B.
GN Names=ECU01.0240;
OS Encephalitozoon cuniculi GB-M1.
OC Eukaryota; Fungi; Microsporidia; Unikaryonidae; Encephalitozoon.
OX NCBI_TaxID=284813;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RC MEDLINE=21576510; PubMed=11719806; DOI=10.1038/35106579;
RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Thonarat F.,
RA Prentier G., Barbe V., Peyrecaillade E., Brottier P., Wincker P.,
RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
RA Weissenbach J., Vivares C.P.;
RT "Genome sequence and gene compaction of the eukaryote parasite
RT Encephalitozoon cuniculi.";
RL Nature 414:450-453(2001).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RC Genoscope;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL391737; CAD24894.1; -.
DR HSSP; P05062; 1005.
DR GO; GO:0004332; F:fructose-bisphosphate aldolase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR000741; Aldolase I.
DR Pfam; PF00274; Glycolytic; 1.
SQ SEQUENCE 338 AA; 37916 MW; 2158A1DAE05A7AF9 CRC64;

Query Match 40.4%; Score 44; DB 2; Length 338;
Best Local Similarity 40.0%; Pred. No. 87;
Matches 8; Conservative 7; Mismatches 3; Indels 2; Gaps 1;

Qy 1 EGYIQNYRCRGDDSKVQEAR 20
Db 299 DGVLETW--RGDDSNIEEAQ 316

RESULT 49
DIAC_HUMAN ID DIAC_HUMAN STANDARD; PRT; 385 AA.
AC Q01459;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Di-N-acetylchitinobiase precursor (EC 3.2.1.-).
GN Name=CTBS; Synonyms=CTB;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RC MEDLINE=92406917; PubMed=1527079;
RA Fisher K.J., Aronson N.N. Jr.;
RT "cloning and expression of the cDNA sequence encoding the lysosomal
RT glycosidase di-N-acetylchitinobiase.";
```

```

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=mx FVB/N; TISSUE=Mammary tumor;
RA Strausberg R.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Adds a myristoyl group to the N-terminal glycine residue
CC of certain cellular proteins (By similarity).
CC -1- CATALYTIC ACTIVITY: Tetradecanoyl-CoA + glycyl-peptide = CoA + N-
CC tetradecanoylglycyl-peptide.
CC -1- SIMILARITY: Belongs to the NMT family.
DR EMBL; BC037647; AAH37647.1; -.
DR HSSP; P30418; 1NMT.
DR MGD; MGI:1202298; Nmt2.
DR GO; GO:0005737; C:cytoplasm; TAS.
DR GO; GO:0004379; F:glycylpeptide N-tetradecanoyltransferase ac. . .; TAS.
DR GO; GO:0006499; P:N-terminal protein myristoylation; TAS.
DR InterPro; IPR000903; Myristoyl_trans.
DR Pfam; PF01233; NMT; 1.
DR Pfam; PF02799; NMT_C; 1.
DR PROSITE; PS00975; NMT_1; 1.
DR PROSITE; PS00976; NMT_2; UNKNOWN_1.
KW Acyltransferase; Transferase.
SQ * SEQUENCE 485 AA; 56537 MW; E4081964FCB3E56E CRC64;

Query Match 40.4%; Score 44; DB 2; Length 485;
Best Local Similarity 53.8%; Pred. No. 1.3e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 YIQNYRCRGDQSK 15
Db 466 YLYNWCPCGTDS 478

```

Search completed: September 7, 2005, 19:57:12
Job time : 86.2326 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2005, 19:49:10 ; Search time 15.3488 Seconds
(without alignments)
125.373 Million cell updates/sec

Title: US-10-812-238B-2

Perfect score: 109

Sequence: 1 EGYIQNYRCRGDDSKVQEAR 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database :

1: PIR 79:*

2: PIR1:*

3: PIR2:*

4: PIR3:*

5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	65	59.6	325	S66668	hydrogen peroxide-
2	52	47.7	504	C64398	hypothetical prote
3	51	46.8	303	T29321	hypothetical prote
4	46.5	42.7	705	A70669	probable acid-CoA
5	46	42.2	518	T19133	hypothetical prote
6	46	42.2	564	T19132	hypothetical prote
7	46	42.2	873	T19131	hypothetical prote
8	45.5	41.7	179	T22521	hypothetical prote
9	45	41.3	607	S77092	hypothetical prote
10	44	40.4	385	A44102	di-N-acetylchitobi
11	43	39.4	274	T32736	hypothetical prote
12	43	39.4	319	AE2823	conserved hypothet
13	43	39.4	341	T16951	hypothetical prote
14	43	39.4	369	D97601	hypothetical prote
15	43	39.4	505	A55923	KNR4 protein - yea
16	43	39.4	542	T10245	glutamyl-tRNA redu
17	43	39.4	565	G89813	DNA polymerase III
18	43	39.4	1150	T18303	DNA binding regula
19	43	39.4	1331	S05011	calcium channel al
20	43	39.4	2166	S11339	calcium channel pr
21	43	39.4	2171	S05054	calcium channel al
22	42	38.5	130	AH0027	30S ribosomal prot
23	42	38.5	182	S28311	hypothetical prote
24	42	38.5	513	B87484	anthranilate synth
25	42	38.5	731	T29129	probable ATP/GTP b
26	42	38.5	957	A84089	hypothetical prote
27	42	38.5	1825	C88400	protein H19M22.1 [
28	42	38.5	1825	T32828	hypothetical prote
29	42	38.5	1994	D86452	protein F6N18.13 [

30	41.5	38.1	325	2	B81364	probable periplasm
31	41	37.6	130	2	C84989	30S ribosomal prot
32	41	37.6	238	2	T11718	hypothetical prote
33	41	37.6	721	2	H82528	L-ascorbate oxidase
34	41	37.6	752	1	PLIVBC	RNA-directed RNA p
35	41	37.6	752	1	PLIVBL	RNA-directed RNA p
36	41	37.6	752	1	PLIVBW	RNA-directed RNA p
37	40.5	37.2	302	2	H83159	conserved hypothet
38	40.5	37.2	373	2	FW0042	activin - fruit fl
39	40.5	37.2	507	2	AD3382	ABC transporter-as
40	40	36.7	132	2	E71356	probable ribosomal
41	40	36.7	150	2	G83840	ferric uptake regu
42	40	36.7	252	2	A48904	aminoglycoside ade
43	40	36.7	290	2	E85529	transcription regu
44	40	36.7	315	2	A90679	mhp operon transcr
45	40	36.7	315	2	B64762	nicotinate phospho
46	40	36.7	400	1	A39130	nicotinate phospho
47	40	36.7	400	1	JQ0756	nicotinate phospho
48	40	36.7	400	2	A06617	nicotinate phospho
49	40	36.7	400	2	F90755	nicotinate phospho
50	40	36.7	400	2	D85619	polysialacturonase
51	40	36.7	422	2	S34199	hypothetical prote
52	40	36.7	435	2	T32195	glutamyl-tRNA redu
53	40	36.7	465	2	T05734	hypothetical prote
54	40	36.7	471	2	S03110	hypothetical prote
55	40	36.7	528	2	T05732	probable glutamyl-
56	40	36.7	530	2	S65773	glutamyl-tRNA redu
57	40	36.7	530	2	G86233	hypothetical prote
58	40	36.7	577	2	T18646	hypothetical prote
59	40	36.7	587	2	S12805	envelysin (EC 3.4.
60	40	36.7	587	2	S41409	envelysin (EC 3.4.
61	40	36.7	644	2	B72559	probable arginyl-t
62	40	36.7	730	2	JH0798	fasciclin IV precu
63	40	36.7	757	1	A60008	RNA-directed RNA p
64	40	36.7	757	1	B60011	RNA-directed RNA p
65	40	36.7	757	1	PLIV61	RNA-directed RNA p
66	40	36.7	757	1	PLIV33	RNA-directed RNA p
67	40	36.7	757	1	PLIV34	RNA-directed RNA p
68	40	36.7	757	1	PLIV68	RNA-directed RNA p
69	40	36.7	757	2	S06212	RNA-directed RNA p
70	40	36.7	924	2	S34926	hypothetical prote
71	40	36.7	2212	2	T28157	erythrocyte membra
72	40	36.7	2238	1	RRVUBY	genome polyprotein
73	40	36.7	2847	2	T28161	hypothetical prote
74	39.5	36.2	348	2	T16794	hypothetical prote
75	39	35.8	168	2	C83505	probable outer mem
76	39	35.8	185	2	H87107	ribosome recycling
77	39	35.8	191	2	A82916	conserved hypothet
78	39	35.8	235	2	B97691	phbF protein (AF02
79	39	35.8	240	2	S65218	hypothetical prote
80	39	35.8	253	2	H87134	3-oxoacyl-[ACP] re
81	39	35.8	280	2	T43011	suppressor protein
82	39	35.8	288	2	E97864	hypothetical prote
83	39	35.8	297	2	C97192	uncharacterized pr
84	39	35.8	301	2	B64189	glycine cleavage s
85	39	35.8	344	2	T05437	hypothetical prote
86	39	35.8	348	2	H85256	hypothetical prote
87	39	35.8	376	2	T32837	hypothetical prote
88	39	35.8	393	2	A75079	hypothetical prote
89	39	35.8	394	2	S77272	hypothetical prote
90	39	35.8	430	2	S75313	serine-tRNA ligase
91	39	35.8	462	2	AH1130	glutamate decarbox
92	39	35.8	470	2	E72376	hypothetical prote
93	39	35.8	521	2	T18896	glypican 1 precurs
94	39	35.8	543	2	E96616	hypothetical prote
95	39	35.8	552	2	T10185	glutamyl-tRNA redu
96	39	35.8	653	2	G96675	hypothetical prote
97	39	35.8	654	2	T24168	hypothetical prote
98	39	35.8	662	2	T41442	omnipotent nonsens
99	39	35.8	662	2	T51948	omnipotent nonsens
100	39	35.8	662	2	I37892	IL12 receptor comp
101	39	35.8	776	2	T09938	hypothetical prote
102	39	35.8	801	2	A99862	Na+/H+ antiporter

103	39	35.8	917	2	I48950	telencephalin prec	176	37	33.9	92	2	G90850	hypothetical prote
104	39	35.8	1162	1	T18199	reverse transcript	177	37	33.9	130	1	R3EC8	ribosomal protein
105	39	35.8	1267	1	MXR31	lambda 3 protein -	178	37	33.9	130	2	AG1007	30S ribosomal chai
106	39	35.8	1267	1	MXR32	lambda 3 protein -	179	37	33.9	130	2	C31150	30S ribosomal subu
107	39	35.8	1267	1	MXR33	lambda 3 protein -	180	37	33.9	130	2	G85995	30S ribosomal subu
108	39	35.8	1268	1	T31420	C-terminal domain-	181	37	33.9	153	2	B83365	transcription regu
109	39	35.8	1403	2	S64142	hypothetical prote	182	37	33.9	158	2	AB0426	transcription elon
110	39	35.8	1603	2	T24098	hypothetical prote	183	37	33.9	159	2	A82219	transcription regu
111	39	35.8	2204	1	R8NZNV	genome polyprotein	184	37	33.9	219	2	T38129	c-8 sterol isomera
112	39	35.8	2493	1	S45734	probable membrane	185	37	33.9	246	2	I50209	reverse transcript
113	39	35.8	3228	2	T21381	hypothetical prote	186	37	33.9	249	2	AB0537	probable membrane
114	39	35.8	9376	2	T14593	syngomycin synth	187	37	33.9	257	2	AH0859	periplasmic fibrin
115	38.5	35.3	448	2	S39348	26S ATP/ubiquitin	188	37	33.9	281	2	S44052	structural protein
116	38.5	35.3	489	2	C87480	conserved hypotet	189	37	33.9	324	2	H86244	lysophospholipase
117	38.5	35.3	662	2	T18233	probable transcript	190	37	33.9	339	2	AH2127	transcription regu
118	38.5	35.3	707	2	F86925	probable acyl-CoA	191	37	33.9	352	2	C84603	probable pectinest
119	38.5	35.3	1287	2	T42658	hypothetical prote	192	37	33.9	359	2	AH3465	oxidoreductase (BC
120	38	34.9	97	2	A82158	hypothetical prote	193	37	33.9	400	2	S46358	protein kinase (BC
121	38	34.9	129	2	A89802	hypothetical prote	194	37	33.9	402	1	F48552	protein kinase (BC
122	38	34.9	160	2	A89809	hypothetical prote	195	37	33.9	434	2	F69081	sensory transducti
123	38	34.9	162	2	F82084	probable 2-demethy	196	37	33.9	435	2	A46206	voltage-gated sodi
124	38	34.9	172	2	F81439	translation initia	197	37	33.9	443	2	H87911	protein B0205.4 [i
125	38	34.9	203	2	S64863	probable membrane	198	37	33.9	450	2	H81338	probable sugar tra
126	38	34.9	250	2	AC0630	response regulator	199	37	33.9	452	2	G90072	hypothetical prote
127	38	34.9	248	2	F64012	vacJ lipoprotein h	200	37	33.9	486	2	F71872	replicative DNA he
128	38	34.9	258	2	T21959	hypothetical prote	201	37	33.9	488	2	B64690	hypothetical prote
129	38	34.9	303	2	G95169	thioredoxin reduct	202	37	33.9	497	2	B97555	replicative DNA he
130	38	34.9	303	2	G98035	thioredoxin-disulf	203	37	33.9	497	2	A02775	hypothetical prote
131	38	34.9	326	2	A97259	mannose-6 phosphate	204	37	33.9	502	2	S38139	fumurate hydratase
132	38	34.9	336	2	T31249	hypothetical prote	205	37	33.9	509	2	S18872	lasi protein - yea
133	38	34.9	345	2	B83260	probable transcript	206	37	33.9	514	2	T00934	legumin-like stora
134	38	34.9	346	2	T49085	hypothetical prote	207	37	33.9	559	2	T48041	probable cytochrom
135	38	34.9	389	2	T27085	hypothetical prote	208	37	33.9	562	2	AF0852	hypothetical prote
136	38	34.9	405	2	S42371	hypothetical prote	209	37	33.9	563	2	S54420	secretory protein i
137	38	34.9	408	2	AF2118	hypothetical prote	210	37	33.9	712	2	A47718	invasion protein i
138	38	34.9	423	2	T21960	hypothetical prote	211	37	33.9	723	2	E82353	reverse transcript
139	38	34.9	423	2	E69052	conserved hypotet	212	37	33.9	757	2	T51428	DNA helicase II VC
140	38	34.9	432	2	C85087	hypothetical prote	213	37	33.9	796	2	E88029	hypothetical prote
141	38	34.9	449	2	F83627	hypothetical prote	214	37	33.9	824	2	T10615	protein F46F5.11 [
142	38	34.9	455	2	E83332	toluate 1,2-dioxyg	215	37	33.9	918	2	E82486	hypothetical prote
143	38	34.9	482	2	H71400	hypothetical prote	216	37	33.9	985	2	B84489	proteinase VCA0223
144	38	34.9	508	2	T474848	neopullulanase - S	217	37	33.9	988	2	F86316	hypothetical prote
145	38	34.9	511	2	G90455	acid phosphatase,	218	37	33.9	1030	2	JC6027	protein T10022.13
146	38	34.9	516	2	S14694	hypothetical prote	219	37	33.9	1038	2	UC6027	115K outer membran
147	38	34.9	533	2	T13607	hypothetical prote	220	37	33.9	1237	2	T14633	hypothetical prote
148	38	34.9	544	2	C86405	probable sphingosi	221	37	33.9	1263	2	T15496	hypothetical prote
149	38	34.9	563	2	AG2923	conserved hypotet	222	37	33.9	1341	2	T18301	latrophilin-2, spl
150	38	34.9	563	2	G97697	hypothetical prote	223	37	33.9	1354	2	T18375	latrophilin-2 (spl
151	38	34.9	572	1	VGNZBA	cell fusion glycop	224	37	33.9	1356	2	T18367	latrophilin-2, spl
152	38	34.9	671	2	AC1407	transketolase homo	225	37	33.9	1369	2	T18379	latrophilin-2 (spl
153	38	34.9	671	2	AC1783	transketolase homo	226	37	33.9	1384	2	T18366	latrophilin-2, spl
154	38	34.9	698	2	A54796	regulatory protein	227	37	33.9	1397	2	T18377	latrophilin-2 (spl
155	38	34.9	718	1	A59084	DNA topoisomerase	228	37	33.9	1399	2	T18370	latrophilin-2 (spl
156	38	34.9	837	2	A42112	mucin-like peptide	229	37	33.9	1407	2	T18381	latrophilin-2 (spl
157	38	34.9	883	2	T49781	related to mutanas	230	37	33.9	1412	2	T18380	latrophilin-2 (spl
158	38	34.9	1071	2	E85343	hypothetical prote	231	37	33.9	1420	2	T18385	latrophilin-2 (spl
159	38	34.9	1071	2	T51224	related to small s	232	37	33.9	1420	2	T17158	CL2AB protein - ra
160	38	34.9	1289	2	AB2217	hypothetical prote	233	37	33.9	1422	2	T18383	latrophilin-2, spl
161	38	34.9	1494	2	T13798	hypothetical prote	234	37	33.9	1435	2	T18387	latrophilin-2 (spl
162	38	34.9	1520	2	A81731	polymorphic membra	235	37	33.9	1435	2	T46611	CL2AB protein - ra
163	38	34.9	1763	2	T19183	hypothetical prote	236	37	33.9	1450	2	T18382	latrophilin-2 (spl
164	38	34.9	1769	2	T19184	hypothetical prote	237	37	33.9	1452	2	T17157	CL2AA protein - ra
165	38	34.9	1812	2	I49350	breast/ovarian can	238	37	33.9	1463	2	T18386	latrophilin-2 (spl
166	38	34.9	4957	2	T03455	ALR protein - huma	239	37	33.9	1463	2	T17159	CL2AC protein - ra
167	38	34.9	5262	2	T03454	ALR protein - huma	240	37	33.9	1465	2	T17159	latrophilin-2 (spl
168	37.5	34.4	255	2	D86752	prophage p12 prote	241	37	33.9	1467	2	T17160	CL2BA protein - ra
169	37.5	34.4	269	2	AC0727	rRNA (guanine-N1-)	242	37	33.9	1478	2	T18388	latrophilin-2 (spl
170	37.5	34.4	277	1	MN1H32	32K nonstructural	243	37	33.9	1478	2	T17185	CL2BC protein - ra
171	37.5	34.4	421	2	C95880	probable protein [244	37	33.9	1487	2	T14324	alpha-latrototoxin r
172	37.5	34.4	421	2	A83139	hypothetical prote	245	37	33.9	1552	2	T05408	hypothetical prote
173	37.5	34.4	503	2	T46852	anthranilate synth	246	37	33.9	1658	2	T42642	phosphoinositide 3
174	37.5	34.4	888	2	H86085	protein T11F1.8 [i	247	37	33.9	1694	2	JQ1896	hypothetical 191.1
175	37.5	34.4	2153	2	T14893	scavenger receptor	248	37	33.9	1984	2	A44396	p-type cation tran

249	37	33.9	2523	2	T18477	hypothetical prote	322	36	33.0	484	2	E84254	hypothetical prote
250	36.5	33.5	174	2	A10618	conserved hypotet	323	36	33.0	487	2	AD3643	adhesin aidA-1 [lm
251	36.5	33.0	269	2	S49049	probable S-adenosy	324	36	33.0	488	1	EXHU	coagulation factor
252	36.5	33.5	269	2	D90945	probable enzyme [i	325	36	33.0	491	2	E87452	ankyrin-related pr
253	36.5	33.5	269	2	G85793	probable enzyme ye	326	36	33.0	527	2	T22867	hypothetical prote
254	36.5	33.5	351	1	SAVZVV	surface antigen pr	327	36	33.0	535	2	T04402	probable glutamyl-
255	36.5	33.5	351	1	SAVZWR	surface antigen pr	328	36	33.0	540	2	S76869	hypothetical prote
256	36.5	33.5	353	1	SAVZVC	surface antigen pr	329	36	33.0	554	2	B85072	hypothetical prote
257	36.5	33.5	354	2	T246877	B20R protein - var	330	36	33.0	574	1	VGNZAR	cell fusion glycop
258	36.5	33.5	354	2	T246877	hypothetical prote	331	36	33.0	574	1	VGNZAR	cell fusion glycop
259	36.5	33.5	355	2	D72104	D9R protein - vari	332	36	33.0	574	1	VGNZHB	cell fusion glycop
260	36.5	33.5	375	2	S62606	semiclotin - mous	333	36	33.0	574	1	VGNZL	cell fusion glycop
261	36.5	33.5	427	2	T14237	hypothetical prote	334	36	33.0	574	2	B28929	cell fusion glycop
262	36.5	33.5	489	2	S66564	acetyl CoA carboxy	335	36	33.0	574	2	S37254	probable membrane
263	36.5	33.5	613	2	T24662	hypothetical prote	336	36	33.0	587	2	C71433	hypothetical prote
264	36.5	33.5	638	2	T24661	hypothetical prote	337	36	33.0	601	2	T21329	probable phosphoe
265	36.5	33.5	878	2	F64425	valine-tRNA ligase	338	36	33.0	602	1	S46676	probable phosphoe
266	36.5	33.5	1022	2	T24663	hypothetical prote	339	36	33.0	660	2	T20569	hypothetical prote
267	36.5	33.5	1143	2	S46122	SNF2 protein homol	340	36	33.0	683	2	S85002	acetate-CoA ligase
268	36.5	33.5	1339	2	JCS508	DNA-directed DNA p	341	36	33.0	705	1	I39759	xylan 1,4-beta-xy
269	36	33.0	36	2	A81164	hypothetical prote	342	36	33.0	710	1	S70965	serine/threonine-s
270	36	33.0	84	1	WMBPWB	gene p protein - p	343	36	33.0	733	2	T04070	hypothetical prote
271	36	33.0	87	2	JQ1683	UL49A protein - hu	344	36	33.0	740	2	AC0119	hypothetical prote
272	36	33.0	98	2	F70769	hypothetical prote	345	36	33.0	742	2	E84285	hypothetical prote
273	36	33.0	111	2	C71401	hypothetical prote	346	36	33.0	789	2	T51310	RNA helicase RH28
274	36	33.0	115	2	F86666	hypothetical prote	347	36	33.0	805	2	S68441	leptin receptor, s
275	36	33.0	132	2	F82916	ribosomal protein	348	36	33.0	808	2	H64474	hypothetical prote
276	36	33.0	138	2	C90460	hypothetical prote	349	36	33.0	819	2	AC2029	hypothetical prote
277	36	33.0	141	2	C64218	hypothetical prote	350	36	33.0	822	2	S77112	DNA mismatch repai
278	36	33.0	142	2	JH0586	ribosomal protein	351	36	33.0	880	2	D69427	conserved hypotet
279	36	33.0	160	2	G95271	hypothetical prote	352	36	33.0	881	1	W2B855	gene 55 protein -
280	36	33.0	161	2	I48158	coagulation factor	353	36	33.0	892	2	S68439	leptin receptor, s
281	36	33.0	168	2	S51484	hypothetical prote	354	36	33.0	894	2	S68437	leptin receptor, s
282	36	33.0	175	2	S61599	probable membrane	355	36	33.0	895	2	S74225	leptin receptor, i
283	36	33.0	180	2	AC3270	hypothetical expor	356	36	33.0	900	2	S68440	leptin receptor, s
284	36	33.0	200	2	D90133	26S proteasome SU	357	36	33.0	1035	2	T42093	phospholipase D [E
285	36	33.0	200	2	T02826	ribosomal protein	358	36	33.0	1036	2	T13732	phospholipase D [E
286	36	33.0	221	2	T04238	hypothetical prote	359	36	33.0	1037	2	T13943	phospholipase D [E
287	36	33.0	236	2	T02314	hypothetical prote	360	36	33.0	1074	2	T17203	phospholipase (EC
288	36	33.0	240	2	E86418	probable eukaryoti	361	36	33.0	1074	2	T13725	phospholipase D [E
289	36	33.0	248	2	F64908	probable dehydroge	362	36	33.0	1075	2	T46635	phospholipase D [E
290	36	33.0	248	2	D90897	probable oxidoredu	363	36	33.0	1131	2	T09701	phytochrome - Scot
291	36	33.0	248	2	C85720	probable oxidoredu	364	36	33.0	1162	2	PC4184	leptin receptor, O
292	36	33.0	259	2	T22862	hypothetical prote	365	36	33.0	1162	2	S68438	leptin receptor, s
293	36	33.0	264	2	F90051	hypothetical prote	366	36	33.0	1191	2	B97116	chromosome segrega
294	36	33.0	282	2	H71823	biotin synthetase	367	36	33.0	1341	2	S09579	tail fiber protein
295	36	33.0	282	2	F64695	biotin synthetase	368	36	33.0	1436	2	A46496	antigen WC1.1 prec
296	36	33.0	286	2	G98209	3-oxoadipate enol-	369	36	33.0	2042	2	T18399	variant-specific s
297	36	33.0	286	2	AB3077	hydrolase [impor	370	36	33.0	3020	2	A43932	mucin 2 precursor,
298	36	33.0	287	2	D81295	cytochrome-c oxida	371	35.5	32.6	80	2	T14387	metallothionein-I
299	36	33.0	295	2	G81785	hypothetical prote	372	35.5	32.6	162	2	F83407	probable phenazine
300	36	33.0	296	2	A48496	transcription regu	373	35.5	32.6	191	2	F86357	Subunit 6b of cyto
301	36	33.0	312	2	E83624	probable transcrip	374	35.5	32.6	242	1	WJMSX3	homeotic protein H
302	36	33.0	329	2	T02546	hypothetical prote	375	35.5	32.6	258	2	G64992	hypothetical prote
303	36	33.0	336	2	T01977	hypothetical prote	376	35.5	32.6	258	2	D91017	hypothetical prote
304	36	33.0	339	2	G90760	probable integrase	377	35.5	32.6	258	2	F85861	hypothetical prote
305	36	33.0	339	2	E85624	probable integrase	378	35.5	32.6	269	2	A46506	leukocyte activati
306	36	33.0	347	2	T35013	probable membrane	379	35.5	32.6	320	2	S58734	probable membrane
307	36	33.0	350	2	D83102	xenobiotic reducta	380	35.5	32.6	363	2	T51341	RNA helicase RH9 [
308	36	33.0	361	2	AF3560	aspartate-semialde	381	35.5	32.6	392	2	H71520	probable hth trans
309	36	33.0	362	2	I40304	outer membrane lip	382	35.5	32.6	411	2	I51285	hepatocyte growth
310	36	33.0	373	2	T35452	probable solute bl	383	35.5	32.6	469	2	T21672	hypothetical prote
311	36	33.0	392	2	D91185	probable regulator	384	35.5	32.6	505	1	TVHUHC	protein-tyrosine k
312	36	33.0	392	2	E86799	prophage pi3 prote	385	35.5	32.6	579	2	AH3225	amidohydrolase [m
313	36	33.0	392	2	D84130	hypothetical prote	386	35.5	32.6	598	2	T40676	hypothetical prote
314	36	33.0	392	2	B86032	probable regulator	387	35	32.1	52	2	S19316	chlorophyll a/b-bi
315	36	33.0	392	2	S47790	xylose operon regu	388	35	32.1	93	2	AF0480	hypothetical prote
316	36	33.0	411	2	H72494	hypothetical prote	389	35	32.1	93	2	S49825	hypothetical prote
317	36	33.0	420	2	T02363	hypothetical prote	390	35	32.1	110	2	T46071	hypothetical prote
318	36	33.0	453	2	T24127	probable chitinase	391	35	32.1	122	2	T04118	mitochondrial proc
319	36	33.0	458	2	S53936	hypothetical prote	392	35	32.1	130	2	JC2280	ribosomal protein
320	36	33.0	461	1	E69360	4-hydroxyphenylace	393	35	32.1	132	2	T45380	ribosomal protein
321	36	33.0	482	1	EXRT	coagulation factor	394	35	32.1	132	2	B85092	hypothetical prote

395 35 32.1 133 2 S29885 ribosomal protein
 396 35 32.1 137 2 T33210 hypothetical prote
 397 35 32.1 154 2 P87526 apag protein limo
 398 35 32.1 160 2 E82382 transcription regu
 399 35 32.1 185 2 T35471 ribosome recycling
 400 35 32.1 187 2 A10520 AmpD protein (anhy
 401 35 32.1 188 2 H95211 anthranilate synth
 402 35 32.1 188 2 B98076 anthranilate synth
 403 35 32.1 189 2 S41410 insecticynanin prec
 404 35 32.1 195 2 S38667 ribonuclease (EC 3
 405 35 32.1 205 2 S34833 stylar protein (al
 406 35 32.1 218 2 T08737 hypothetical prote
 407 35 32.1 231 2 AD2892 conserved hypothet
 408 35 32.1 239 1 R3223 ribosomal protein
 409 35 32.1 240 2 T50831 superoxide dismuta
 410 35 32.1 240 2 C83702 transcription regu
 411 35 32.1 250 2 C84637 probable replicati
 412 35 32.1 252 2 T46247 hypothetical prote
 413 35 32.1 264 2 T31238 trak protein homol
 414 35 32.1 272 2 T15351 hypothetical prote
 415 35 32.1 281 1 B47629 cell surface glyco
 416 35 32.1 285 2 A86419 probable eukaryoti
 417 35 32.1 287 2 T25064 hypothetical prote
 418 35 32.1 287 2 B82806 integral membrane
 419 35 32.1 289 2 A95158 homoserine kinase
 420 35 32.1 289 2 A98024 homoserine kinase
 421 35 32.1 289 2 T16530 hypothetical prote
 422 35 32.1 291 2 T28846 hypothetical prote
 423 35 32.1 291 2 AE2000 glutamyl-tRNA synt
 424 35 32.1 295 2 C81210 probable transport
 425 35 32.1 312 2 T35940 36K protein - appl
 426 35 32.1 320 2 B44059 hypothetical prote
 427 35 32.1 327 2 G97667 protein B0361.5 [i
 428 35 32.1 333 2 A88504 NCA3 protein precu
 429 35 32.1 337 2 S55932 hypothetical prote
 430 35 32.1 338 2 E71195 protein kinase-like
 431 35 32.1 342 2 T45577 hypothetical prote
 432 35 32.1 347 2 B72596 probable ATP-bindi
 433 35 32.1 348 2 C97633 hypothetical prote
 434 35 32.1 348 2 AF2856 probable aldose 1-
 435 35 32.1 356 2 H72395 gene wnt8 protein
 436 35 32.1 361 2 I50505 hypothetical prote
 437 35 32.1 361 2 A68641 hieB bifunctional
 438 35 32.1 362 2 F64070 casein kinase-1 ho
 439 35 32.1 365 2 S46357 hypothetical prote
 440 35 32.1 377 2 T33765 hypothetical prote
 441 35 32.1 377 2 T47439 elastase inhibitor
 442 35 32.1 379 2 S27383 hypothetical prote
 443 35 32.1 380 2 T32163 histidinol-phospha
 444 35 32.1 384 1 S41584 histidinol-phospha
 445 35 32.1 388 2 A39756 circumporozone p
 446 35 32.1 389 1 A48329 circumporozone p
 447 35 32.1 392 2 A86252 hypothetical prote
 448 35 32.1 395 2 E86152 chalcone synthase
 449 35 32.1 402 2 E69107 hypothetical prote
 450 35 32.1 404 2 S13648 SCU1 protein - yea
 451 35 32.1 405 2 S05428 circumporozone p
 452 35 32.1 412 1 OZZQAP circumporozone p
 453 35 32.1 416 2 C81376 hypothetical prote
 454 35 32.1 417 2 T33376 hypothetical prote
 455 35 32.1 424 2 A54533 circumporozone p
 456 35 32.1 428 2 S36856 glycolipid 2-alpha
 457 35 32.1 442 2 A54529 circumporozone p
 458 35 32.1 442 2 JC5594 jerky gene protein
 459 35 32.1 444 2 S57989 probable membrane
 460 35 32.1 448 2 A43304 phosphomannomutase
 461 35 32.1 454 2 A41025 DNA helicase [limo
 462 35 32.1 462 2 S33084 gamma-tubulin - Eu
 463 35 32.1 462 2 H95914 hypothetical membr
 464 35 32.1 465 2 F75524 hypothetical prote
 465 35 32.1 476 2 T23083 3-isopropylmalate
 466 35 32.1 503 2 S38005 hypothetical prote
 467 35 32.1 504 2 A81922 probable cell-surf

468 35 32.1 505 2 F95345 FixL Oxygen regula
 469 35 32.1 505 2 S39984 nitrogen fixation
 470 35 32.1 509 2 T02864 probable Zn finger
 471 35 32.1 509 2 T37587 hypothetical trp-a
 472 35 32.1 518 2 T09354 hypothetical prote
 473 35 32.1 524 2 A32617 phytoene dehydroge
 474 35 32.1 525 1 X2AD10 penton protein (II
 475 35 32.1 525 2 T25900 hypothetical prote
 476 35 32.1 533 2 B26446 vif2 protein - Ag
 477 35 32.1 539 2 A54294 cell division cont
 478 35 32.1 548 2 T22137 hypothetical prote
 479 35 32.1 559 2 A57474 extracellular matr
 480 35 32.1 583 2 S46265 cysteine proteinas
 481 35 32.1 598 2 S51456 probable membrane
 482 35 32.1 608 1 S28313 hypothetical prote
 483 35 32.1 617 2 S08317 hydrogenase (EC 1.
 484 35 32.1 618 1 B43255 hydrogenase (EC 1.
 485 35 32.1 618 2 S09251 hydrogenase (EC 1.
 486 35 32.1 619 1 JH0776 hydrogenase (EC 1.
 487 35 32.1 619 2 A45625 phosphoenolpyruvat
 488 35 32.1 627 2 E86190 hypothetical prote
 489 35 32.1 651 2 T32875 hypothetical prote
 490 35 32.1 655 2 T34219 serine/threonine-s
 491 35 32.1 656 2 T10664 hypothetical prote
 492 35 32.1 670 2 T09205 hypothetical prote
 493 35 32.1 670 2 T09274 hypothetical prote
 494 35 32.1 678 2 T04832 probable serine/th
 495 35 32.1 729 2 T23972 hypothetical prote
 496 35 32.1 754 2 C96807 hypothetical prote
 497 35 32.1 759 2 T43031 DNA topoisomerase
 498 35 32.1 762 2 G88436 protein T04A8.13 [i
 499 35 32.1 773 2 I59351 carnitine O-palmit
 500 35 32.1 787 2 H86707 formate C-acetyltr

ALIGNMENTS

RESULT 1

S66668 hydrogen peroxide-inducible protein - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004

C:Accession: S66668

R: Egawa, K.; Yoshiwara, M.; Shibamura, M.; Nose, K.

FEBS Lett. 372: 74-77, 1995

A:Title: Isolation of a novel ras-recision gene that is induced by hydrogen peroxide from

A:Reference number: S66668; MUID:96032549; PMID:7556647

A:Accession: S66668

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-325 <EGL>

A:Cross-references: UNIPROT:Q61469; EMBL:L43371; NID:gl161099; PIDN:AAA85353.1; PID:gl161

A>Note: in the authors' translation residues 241-260 do not match the nucleotide sequence

Query Match 59.6%; Score 65; DB 2; Length 325;

Best Local Similarity 50.0%; Pred No. 0.0061;

Matches 10; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 EGYIQNYRCRGDSKVQEAR 20

:|||:|:|:|:|:|:|

Db 256 DGYIEDYICQNGEKVKEGR 275

RESULT 2

C64398 hypothetical protein homolog MJ0787 - Methanococcus jannaschii

C:Species: Methanococcus jannaschii

C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004

C:Accession: C64398

R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, C.

; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodex, A.;

rsen, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.

Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodowsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.

A:Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*

A:Reference number: A64300; MUID:96337999; PMID:8688087

A:Accession: C64398

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-504 <BUL>

A:Cross-references: UNIPROT:Q58197; GB:U67523; GB:L77117; NID:g2826319; PIDN:AAB98783.1.7

C:Genetics:

A:Map position: FOR710775-712289

Query Match 47.7%; Score 52; DB 2; Length 504;
Best Local Similarity 57.9%; Pred. No. 1.4;
Matches 11; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 EGYIYNYRCRGDDSKVOEA 19
Db 457 EAGINAYPCNGDDKKVLEA 475

RESULT 3

T29321

C:Species: *Caenorhabditis elegans*

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000

C:Accession: T29321

R:Pauley, A.; Gattung, S.

A:Description: The sequence of *C. elegans* cosmid M01E11.

A:Reference number: Z20605

A:Accession: T29321

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-303 <PAU>

A:Cross-references: EMBL:U80450; PIDN:AA837827.1; GSPDB:GN00019; CESP:M01E11.4

A:Experimental source: strain Bristol N2; clone M01E11

C:Genetics:

A:Gene: CESP:M01E11.4

A:Map position: 1

A:Introns: 32/3; 135/2; 152/2; 267/3

C:Superfamily: *Caenorhabditis elegans* hypothetical protein M01E11.4

Query Match 46.8%; Score 51; DB 2; Length 303;
Best Local Similarity 50.0%; Pred. No. 1.2;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 EGYIYNYRCRGDDSKVOE 18
Db 34 EGFVTRGCGRLDSQIOE 51

RESULT 4

A70869

N:Alternate names: hypothetical protein pks002a

C:Species: *Mycobacterium tuberculosis*

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C:Accession: A70669; S73073

R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: A70669

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-705 <COL>

A:Cross-references: UNIPROT:P96283; GB:Z83858; GB:AL123456; NID:g3261675; PIDN:CAB06101.7

A:Experimental source: strain H37Rv

R:Smith, D.R.; Robinson, K.

submitted to the EMBL Data Library, September 1994
A:Description: Mycobacterium tuberculosis cosmid tbc2.
A:Reference number: S73053
A:Accession: S73073
A:Molecule type: DNA
A:Residues: 1-486, 'OQ', 489-705 <SMI>
A:Cross-references: EMBL:U00024; NID:G560506; PIDN:AAAS0930.1; PID:G560509
C:Genetics:
A:Gene: fadD22
C:Keywords: acid-thiol ligase; carrier protein; phosphopantetheine; phosphoprotein
F:51-486/Domain: acetate-CoA ligase homology <ACL>
F:544-615/Domain: acyl carrier protein homology <ACPI>
F:579/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 42.7% Score 46.5; DB 2; Length 705;
Best Local Similarity 50.0%; Pred. No. 16;
Matches 8; Conservative 5; Mismatches 2; Indels 1; Gaps 1;

Qy 1 EGYIYRCRGDSDKV 16
::: ||| |||:
Db 378 DGWV-TYRCADDTTEV 392

RESULT 5
T19132
hypothetical protein C09F9.5b - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T19133
R:Smyle, R.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19078
A:Accession: T19133
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-518 <WIL>
A:Cross-references: EMBL:Z81465; PIDN:CAB03864.1; GSPDB:GN000020; CESP:C09F9.5b
A:Experimental source: clone C09F9
C:Genetics:
A:Gene: CESP:C09F9.5b
A:Map position: 2
A:Introns: 34/2; 69/3; 139/3; 222/3; 284/2; 335/2; 377/1

Query Match 42.2% Score 46; DB 2; Length 518;
Best Local Similarity 50.0%; Pred. No. 14;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 3 YIYRCRGDSDKV 18
::: ||| |||:
Db 51 YMKEYRCFPDESQVYE 66

RESULT 6
T19132
hypothetical protein C09F9.5a - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T19132
R:Smyle, R.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19078
A:Accession: T19132
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-564 <WIL>
A:Cross-references: EMBL:Z81465; PIDN:CAB03863.1; GSPDB:GN000020; CESP:C09F9.5a
A:Experimental source: clone C09F9
C:Genetics:
A:Gene: CESP:C09F9.5a
A:Map position: 2
A:Introns: 34/2; 69/3; 139/3; 222/3; 284/2; 335/2; 377/1; 505/1

Query Match 42.2% Score 46; DB 2; Length 564;

A;Experimental source: strain Bristol N2; clone F54D7

C;Genetics:

A;Gene: CESP:F54D7.4

A;Map position: 1

A;Introns: 10/1; 54/2; 78/3; 124/1; 203/1; 258/2

Query Match 39.4%; Score 43; DB 2; Length 274;

Best Local Similarity 57.1%; Pred. No. 23;

Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 EGYIQNYRCRGDDSK 14

Db 222 EDHIGKRCIGDNS 235

RESULT 12

AE2823 conserved hypothetical protein Atu2011 [imported] - Agrobacterium tumefaciens (strain C58)

C;Species: Agrobacterium tumefaciens

C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004

C;Accession: AE2823

R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutuyavin, T.; Levy, R.; Li, M.; McClell

; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

ster, E.W.

A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A;Reference number: AB2577; MUID:21608550; PMID:11743193

A;Accession: AE2823

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-319 <KUR>

A;Cross-references: UNIPROT:Q8UDV8; GB:AE008688; PIDN:AAL43003.1; PID:g17740465; GSPDB:C

A;Experimental source: strain C58 (Dupont)

C;Genetics:

A;Gene: Atu2011

A;Map position: circular chromosome

Query Match 39.4%; Score 43; DB 2; Length 319;

Best Local Similarity 50.0%; Pred. No. 27;

Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 GYIQNYRCRGDDSK 15

Db 50 GYLQKYLVEGNDNR 63

RESULT 13

T16951

hypothetical protein T28D9.3 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004

C;Accession: T16951

R;Fulton, L.

submitted to the EMBL Data Library, June 1995

A;Description: The sequence of C. elegans cosmid T28D9.

A;Reference number: Z18614

A;Accession: T16951

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-341 <FUL>

A;Cross-references: UNIPROT:Q10022; EMBL:U28738; NID:g861262; PID:g861267; PIDN:AAA68312

A;Experimental source: strain Bristol N2

C;Genetics:

A;Gene: CESP:T28D9.3

A;Introns: 87/3; 127/1; 172/2; 209/3; 269/3; 303/1

Query Match 39.4%; Score 43; DB 2; Length 341;

Best Local Similarity 50.0%; Pred. No. 28;

Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 3 YIQNYRCRGDDSKVQEAR 20

Db 173 YITDYTCGPPELVLEAR 190

RESULT 14

D97601

hypothetical protein AGR_C_3649 [imported] - Agrobacterium tumefaciens (strain C58, Cerec

C;Species: Agrobacterium tumefaciens

C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004

C;Accession: D97601

R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,

A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;

Science 294, 2323-2328, 2001

A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A;Reference number: A97359; MUID:21608551; PMID:11743194

A;Accession: D97601

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-369 <KUR>

A;Cross-references: UNIPROT:Q8UDV8; GB:AE007869; PIDN:AAK87765.1; PID:g15157135; GSPDB:G

C;Genetics:

A;Gene: AGR_C_3649

A;Map position: circular chromosome

Query Match 39.4%; Score 43; DB 2; Length 369;

Best Local Similarity 50.0%; Pred. No. 31;

Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 GYIQNYRCRGDDSK 15

Db 100 GYLQKYLVEGNDNR 113

RESULT 15

AS5923

KNR4 protein - yeast (Saccharomyces cerevisiae)

N;Alternate names: protein 505; protein G8553; protein YGR229c; SM11 protein

C;Species: Saccharomyces cerevisiae

C;Date: 08-Sep-1995 #sequence_revision 08-Sep-1995 #text_change 09-Jul-2004

C;Accession: AS5923; S57693; S64553; S63909

R;Hong, Z.; Mann, P.; Brown, N.H.; Tran, L.E.; Shaw, K.J.; Hare, R.S.; DiDomenico, B.

Mol. Cell. Biol. 14, 1017-1025, 1994

A;Title: Cloning and characterization of KNR4, a yeast gene involved in (1,3)-beta-glucar

A;Reference number: AS5923; MUID:94119053; PMID:8289782

A;Accession: AS5923

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-505 <HON>

A;Cross-references: UNIPROT:P32566; GB:LI13164; NID:g311112; PIDN:AAA67469.1; PID:g311113

A;Note: authors translated the codon AAC for residue 388 as Ala, and GCA for residue 389

R;van der Aart, Q.J.M.; Kleine, K.; Steensma, H.Y.

submitted to the EMBL Data Library, June 1995

A;Description: Sequence analysis of the 43 KB CRM1-YLM9-PET54-SM11-PHO81-YHB4-PFK1 region

A;Reference number: S57680

A;Accession: S57693

A;Molecule type: DNA

A;Residues: 1-505 <VAN>

A;Cross-references: EMBL:X87941; NID:g886908; PID:g886922

A;Experimental source: strain S288C

R;van der Aart, Q.J.M.; Steensma, H.Y.

submitted to the Protein Sequence Database, May 1996

A;Reference number: S64553

A;Accession: S64553

A;Molecule type: DNA

A;Residues: 1-505 <VAN>

A;Cross-references: EMBL:Z73013; NID:g1323411; PID:e243665; PID:g1323413; MIPS:YGR229C

A;Experimental source: strain S288C

R;van der Aart, Q.J.M.; Kleine, K.; Steensma, H.Y.

Yeast 12, 385-390, 1996

A;Title: Sequence analysis of the 43 kb CRM1-YLM9-PET54-DIE2-SM11-PHO81-YHB4-PFK1 region

A;Reference number: S63896; MUID:96267763; PMID:8701610

A;Accession: S63909

A;Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
A:Residues: 1-505 <VAF>
A:Cross-references: EMBL:X87941; NID:g886908; PIDN:CAA61178.1; PID:g886922
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995
C:Genetics:
A:Gene: SGD:SM11; KNR4
A:Cross-references: SGD:S0003461; MIPS:YGR229C
A:Map position: 7R
C:Superfamily: cell wall assembly and cell proliferation coordinating protein, KNR4 type

Query Match 39.4%; Score 43; DB 2; Length 505;
Best Local Similarity 72.7%; Pred. No. 42;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 10 RGDGSKVQEAR 20
| | | | | | | | | |
Db 487 RSDDAKVEAR 497

RESULT 16
T10245
glutamyl-tRNA reductase (EC 1.2.1.-) 2 - cucumber
C:Species: Cucumis sativus (cucumber)
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: T10245
R.Tanaka, R.; Yoshida, K.; Nakayashiki, T.; Masuda, T.; Tsuji, H.; Inokuchi, H.; Tanaka, R.
Plant Physiol. 110, 1223-1230, 1996
A:Title: Differential expression of two hema mRNAs encoding glutamyl-tRNA reductase prot
A:Reference number: Z16983; MUID:97088717; PMID:8934625
A:Accession: T10245
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-542 <TAN>
A:Cross-references: UNIPROT:P49295; EMBL:D67088; NID:g1015318; PIDN:BAAL1091.1; PID:g101
A:Experimental source: cv. Aonagajibai; greening cotyledons
C:Genetics:
A:Gene: hemA2
A:Genome: nuclear
C:Function:
A:Description: catalyzes the reduction of glutamyl-tRNA(Glu) by NADPH to glutamic acid 1
A:Pathway: aminolevulinate biosynthesis; porphyrin biosynthesis
C:Superfamily: glutamyl-tRNA reductase
C:Keywords: aminolevulinate biosynthesis; chloroplast; NADP; oxidoreductase; porphyrin b

Query Match 39.4%; Score 43; DB 2; Length 542;
Best Local Similarity 57.1%; Pred. No. 45;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 GYIQNYRCRGDSK 15
| | | | | | | | | |
Db 491 GPIQLRLCDGNSR 504

RESULT 17
G9813
DNA polymerase III gamma and tau subunits [imported] - Staphylococcus aureus (strain N31
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: G9813
R.Kuroda, M.; Ohca, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: G9813
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-565 <KUR>
A:Cross-references: UNIPROT:Q99WC5; GB:BA000018; PID:g13700368; PIDN:BABA1666.1; GSPDB:G
A:Experimental source: strain N315
C:Genetics:
A:Gene: dnaX

C:Superfamily: DNA-directed DNA polymerase III gamma chain

Query Match 39.4%; Score 43; DB 2; Length 565;
Best Local Similarity 44.4%; Pred. No. 47;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 3 YIQNYRCRGDDSKVQEAR 20
| | | | | | | | | |
Db 525 YIQNRKNEGDDMPKQQAQ 542

RESULT 18
T18303
DNA binding regulatory protein - Emericella nidulans
C:Species: Emericella nidulans, Aspergillus nidulans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18303
R.Murphy, R.L.; Andrianopoulos, A.; Davis, M.A.; Hynes, M.J.
Mol. Microbiol. 23, 591-602, 1997
A:Title: Identification of amdX, a new Cys-2-His-2 (C2H2) zinc-finger gene involved in th
A:Reference number: Z18871; MUID:97197203; PMID:9044292
A:Accession: T18303
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1150 <MUR>
A:Cross-references: UNIPROT:P79045; EMBL:U56100; NID:g1877524; PID:g1877525; PIDN:AAB530C
C:Genetics:
A:Gene: amdX
A:introns: 32/1

Query Match 39.4%; Score 43; DB 2; Length 1150;
Best Local Similarity 52.6%; Pred. No. 96;
Matches 10; Conservative 2; Mismatches 5; Indels 2; Gaps 1;

QY 4 IQNYRCR--GDDSKVQEAR 20
: | | | | | | | | | |
Db 802 VQNSCRSLSPDDSKSDLR 820

RESULT 19
S05011
calcium channel alpha-1 chain, dihydropyridine sensitive, cardiac (clone HTDP 2.0) - rat
N:Alternate names: dihydropyridine receptor, cardiac; voltage-dependent calcium channel t
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 11-Jan-2000
C:Accession: S05011
R.Slash, D.F.; Engle, D.B.; Varadi, G.; Lotan, I.; Singer, D.; Dascal, N.; Schwartz, A.
FEBS Lett. 250, 509-514, 1989
A:Title: Evidence for the existence of a cardiac specific isoform of the alpha(1) subunit
A:Reference number: S05011; MUID:89325623; PMID:2546809
A:Accession: S05011
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-1331 <SLI>
C:Superfamily: voltage-dependent calcium channel protein alpha-1 chain
C:Keywords: cardiac muscle; heart

Query Match 39.4%; Score 43; DB 2; Length 1331;
Best Local Similarity 50.0%; Pred. No. 11e+02;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 5 QNYRCRGDDSKVQEAR 20
: | | | | | | | | | |
Db 1288 ENSPCRGDSSAARRAR 1303

RESULT 20
S11339
calcium channel protein - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S11339
R.Biel, M.; Ruth, P.; Bosse, E.; Hüllin, R.; Stuehmer, W.; Flockerzi, V.; Hofmann, F.

FEBS Lett. 269, 409-412, 1990
A:Title: Primary structure and functional expression of a high voltage activated calcium
A:Reference number: S11339; MUID:90382589; PMID:2169433
A:Accession: S11339
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2166 <BIE>
A:Cross-references: UNIPROT:P15381; GB:X55763; NID:g1507; PIDN:CAA39289.1; PID:g1508
C:Superfamily: voltage-dependent calcium channel protein alpha-1 chain

Query Match 39.4%; Score 43; DB 2; Length 2166;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 5 QNYRCRGDDSKVQEAR 20
: || || || || : ||
Db 2023 ENSPCRCGDSAAARR 2038

RESULT 21
S05054
calcium channel alpha-1 chain, dihydropyridine sensitive, cardiac (clone lambda-CCAR34)
N:Alternate names: dihydropyridine receptor, cardiac
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C:Accession: S05054
R:Mikami, A.; Imoto, K.; Tanabe, T.; Niidome, T.; Mori, Y.; Takeshima, H.; Narumiya, S.;
Nature 340, 230-233, 1989
A:Title: Primary structure and functional expression of the cardiac dihydropyridine-sens
A:Reference number: S05054; MUID:89330539; PMID:2474130
A:Accession: S05054
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-2171 <MIK>
A:Cross-references: UNIPROT:P15381; EMBL:X15539; NID:g1509; PIDN:CAA33546.1; PID:g1510
C:Superfamily: voltage-dependent calcium channel protein alpha-1 chain
C:Keywords: cardiac muscle; heart; transmembrane protein

Query Match 39.4%; Score 43; DB 2; Length 2171;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 5 QNYRCRGDDSKVQEAR 20
: || || || || : ||
Db 2028 ENSPCRCGDSAAARR 2043

RESULT 22
AH0027
30S ribosomal protein S8 [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Accession: AH0027
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
Il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AH0027
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-130 <KUR>
A:Cross-references: UNIPROT:Q8ZJ98; GB:AL590842; PIDN:CAC89083.1; PID:g15978323; GSPDB:G
C:Genetics:
A:Gene: rpsH
C:Superfamily: Escherichia coli ribosomal protein S8

Query Match 38.5%; Score 42; DB 2; Length 130;
Best Local Similarity 38.9%; Pred. No. 16;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 EGYIQNYRCRGDDSKVQ 18

Db 43 EGFIEDFKIEGDTKPVL 60
||:|::: || | |
RESULT 23
S28311
hypothetical protein F02A9.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 09-Jul-2004
C:Accession: S28311
R:Berks, M.
submitted to the EMBL Data Library, December 1992
A:Reference number: S28296
A:Accession: S28311
A:Molecule type: DNA
A:Residues: 1-182 <BER>
A:Cross-references: UNIPROT:P34383; EMBL:Z19555; NID:g6705; PID:g6709
C:Genetics:
A:Introns: 170/1
C:Superfamily: Caenorhabditis elegans hypothetical protein F02A9.3

Query Match 38.5%; Score 42; DB 2; Length 182;
Best Local Similarity 46.7%; Pred. No. 22;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 EGYIQNYRCRGDDSK 15
: || || : ||
Db 140 KGYYAEYKALSDDAK 154

RESULT 24
B87484
anthranilate synthase component I [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: B87484
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.J.;
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Emolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: AB87249; MUID:21173698; PMID:11259647
A:Accession: B87484
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-513 <STO>
A:Cross-references: UNIPROT:Q9A731; GB:AE005673; NID:g13423344; PIDN:AAK23870.1; GSPDB:G
C:Genetics:
A:Gene: CC1895
C:Superfamily: anthranilate synthase component I

Query Match 38.5%; Score 42; DB 2; Length 513;
Best Local Similarity 50.0%; Pred. No. 63;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 7 YRCRGDDSKVQ 18
: || || || : ||
Db 73 WRCRGDQAEIAE 84

RESULT 25
T29129
probable ATP/GTP binding protein - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C:Accession: T29129
R:Parkhill, J.; Barrrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1998
A:Reference number: Z17215
A:Accession: T29129
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-731 <PAR>

A;Cross-references: UNIPROT:O86557; EMBL:AL031350; NID:e1316892; PID:e1316912; PIDN:CAA2
 C;Genetics:
 A;Note: SC1F2.20

Query Match 38.5%; Score 42; DB 2; Length 731;
 Best Local Similarity 43.8%; Pred. No. 90;
 Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 5 QNYRCRGDDSKVQEAR 20
 Db 460 EQYRCGGSDAYIRWAR 475
 : ||| | | : : : ||
 : ||| | | : : : ||

RESULT 26
 A84089
 hypothetical protein BH3513 [imported] - Bacillus halodurans (strain C-125)
 C;Species: Bacillus halodurans
 C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
 C;Accession: A84089
 R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
 Nucleic Acids Res. 28, 4317-4331, 2000
 A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
 A;Reference number: A83650; MUID:20512592; PMID:11059132
 A;Accession: A84089
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-957 <STO>
 A;Cross-references: UNIPROT:Q9K760; GB:AP001519; GB:BA000004; NID:g10176109; PIDN:BAB072
 A;Experimental source: strain C-125
 C;Genetics:
 A;Gene: BH3513

Query Match 38.5%; Score 42; DB 2; Length 957;
 Best Local Similarity 60.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 GYIQNYRCRG 11
 Db 926 GFIKNYGCKG 935
 : ||| | | :
 : ||| | | :

RESULT 27
 C89400
 protein H19M22.1 [imported] - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
 C;Accession: C89400
 R;anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
 A;Reference number: A75000; MUID:99069613; PMID:9851916
 A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
 A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
 A;Accession: C89400
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-1825 <STO>
 A;Cross-references: UNIPROT:Q9UB28; UNIPROT:Q9UB29; GB:chr_III; PIDN:AAB94997.1; PID:g27
 C;Genetics:
 A;Gene: H19M22.1
 A;Map position: 3

Query Match 38.5%; Score 42; DB 2; Length 1825;
 Best Local Similarity 41.2%; Pred. No. 2.2e+02;
 Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 GYIQNYRCRGDDSKVQE 18
 Db 459 GYVDKYKCRYAPSGTQQ 475
 : ||| | | :
 : ||| | | :

RESULT 28
 T32828

hypothetical protein H19M22.1 - Caenorhabditis elegans (fragment)
 C;Species: Caenorhabditis elegans
 C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
 C;Accession: T32828
 R;Wilson, R.; Wamsley, P.
 submitted to the EMBL Data Library, December 1997
 A;Description: The sequence of C. elegans cosmid H19M22.
 A;Reference number: Z21229
 A;Accession: T32828
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-1825 <WIL>
 A;Cross-references: UNIPROT:Q9UB28; UNIPROT:Q9UB29; EMBL:AF040648; PIDN:AAB94997.1; GSPDI
 A;Experimental source: strain Bristol N2; clone H19M22
 C;Genetics:
 A;Gene: CESP:H19M22.1
 A;Map position: 3
 A;Introns: 271/2

Query Match 38.5%; Score 42; DB 2; Length 1825;
 Best Local Similarity 41.2%; Pred. No. 2.2e+02;
 Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 GYIQNYRCRGDDSKVQE 18
 Db 459 GYVDKYKCRYAPSGTQQ 475
 : ||| | | :
 : ||| | | :

RESULT 29
 D86452
 protein F6N18.13 [imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C;Accession: D86452
 R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khavkin, E.; Kim, C.C
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maili, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A;Reference number: A86141; MUID:21016719; PMID:11130712
 A;Accession: D86452
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-1994 <STO>
 A;Cross-references: UNIPROT:Q9LPI9; GB:AE005172; NID:g6714281; PIDN:AAF25977.1; GSPDB:GN
 C;Genetics:
 A;Gene: F6N18.13
 A;Map position: 1

Query Match 38.5%; Score 42; DB 2; Length 1994;
 Best Local Similarity 38.9%; Pred. No. 2.5e+02;
 Matches 7; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 EGYQNYRCRGDDSKVQE 18
 Db 1827 EGYLEDYPPKRNDRLE 1844
 : ||| | | : : : ||
 : ||| | | : : : ||

RESULT 30
 B81364
 probable periplasmic protein Cj0906c [imported] - Campylobacter jejuni (strain NCTC 11161
 C;Species: Campylobacter jejuni
 C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
 C;Accession: B81364
 R;Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling
 C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell
 Nature 403, 663-668, 2000
 A;Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp

A;Reference number: A81250; MUID:20150912; PMID:10688204

A;Accession: B81364

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-325 <PAR>

A;Cross-references: UNIPROT:Q9PP25; GB:AL1139076; GB:AL111168; NID:g9668128; PIDN:CAB7316

A;Experimental source: serotype O2, strain NCTC 11168

C;Genetics:

A;Gene: Cj0906c

C;Superfamily: Helicobacter pylori hypothetical protein jhp0467

Query Match 38.1%; Score 41.5; DB 2; Length 325;

Best Local Similarity 50.0%; Pred. No. 48;

Matches 10; Conservative 1; Mismatches 2; Indels 7; Gaps 1;

QY 3 YIQNYRCRGD-----DSK 15

DB 297 YTKNYRFRGDKILVVKIDSK 316

RESULT 31

CB4989

30S ribosomal protein S8 [imported] - Buchnera sp. (strain APS)

C;Species: Buchnera sp.

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001

C;Accession: CB4989

R;Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.

Nature 407, 81-86, 2000

A;Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A

A;Reference number: A84930; MUID:20445173; PMID:10993077

A;Accession: CB4989

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-130 <STO>

A;Cross-references: GB:AP000398; GSPDB:GN00144

A;Experimental source: strain APS

C;Genetics:

A;Gene: rpsH; BUS10

C;Superfamily: Escherichia coli ribosomal protein S8

Query Match 37.6%; Score 41; DB 2; Length 130;

Best Local Similarity 58.3%; Pred. No. 23;

Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EGYIQNYRCRGD 12

DB 43 EGYIKDYNVTGD 54

RESULT 32

T11718

hypothetical protein SPCC777.12c - fission yeast (Schizosaccharomyces pombe)

C;Species: Schizosaccharomyces pombe

C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004

C;Accession: T11718

R;Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, August 1998

A;Reference number: Z17318

A;Accession: T11718

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-238 <SEE>

A;Cross-references: UNIPROT:O74551; EMBL:AL031532

A;Experimental source: strain 972h(-)

C;Genetics:

A;Map position: IIIR

A;Introns: 93/3

A;Note: SPCC777.12c

C;Superfamily: Schizosaccharomyces hypothetical protein SPCC777.12c

Query Match

Best Local Similarity

Matches

37.6%; Score 41; DB 2; Length 238;

41.2%; Pred. No. 42;

7; Conservative

3; Mismatches

7; Indels

0; Gaps

QY 3 YIQNYRCRGDSSKVOEA 19

DB 16 FVAAYSCEGDSAAIEA 32

RESULT 33

H82528

L-ascorbate oxidase XP2677 [imported] - Xylella fastidiosa (strain 9a5c)

C;Species: Xylella fastidiosa

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C;Accession: H82528

R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequenc

Nature 406, 151-157, 2000

A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A;Reference number: A82515; MUID:20365717; PMID:10910347

A;Note: for a complete list of authors see reference number A59328 below

A;Accession: H82528

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-721 <SIM>

A;Cross-references: UNIPROT:Q9PA43; GB:AE004073; GB:AE003849; NID:g9107904; PIDN:AAF85474

A;Experimental source: strain 9a5c

R;Simpson, A.J.G.; Reinach, F.C.; Armada, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Al

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H.

as-Neco, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Praga, J.S.; Franca, S.C.; Franco, M.C.; Frohme

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigre

chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.

; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Miyaki, C.Y.;

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasaki

A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira

M.; Teuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

A;Reference number: A59328

A;Contents: annotation

C;Genetics:

A;Gene: XF2677

Query Match 37.6%; Score 41; DB 2; Length 721;

Best Local Similarity 42.9%; Pred. No. 1.3e+02;

Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 7 YRCRGDSSKVOEAR 20

DB 282 YACRGNDGKLRKK 295

RESULT 34

PIIVBC

RNA-directed RNA polymerase (EC 2.7.7.48) 1 - influenza B virus (strain B/Ann Arbor/1/66

N;Alternate names: P1 protein; PB1 protein

C;Species: influenza B virus

C;Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 09-Jul-2004

C;Accession: C28604

R;DeBorde, D.C.; Donabedian, A.M.; Herlocher, M.L.; Naeye, C.W.; Maassab, H.F.

Virology 163, 429-443, 1988

A;Title: Sequence comparison of wild-type and cold-adapted B/Ann Arbor/1/66 influenza vi

A;Reference number: A28604; MUID:88179548; PMID:3354202

A;Accession: C28604

A;Molecule type: genomic RNA

A;Residues: 1-752 <DEB>

A;Cross-references: UNIPROT:PI3871

C;Genetics:

A;Map position: segment 2

A;Superfamily: influenza virus RNA-directed RNA polymerase 1

C;Keywords: nucleotidyltransferase

Query Match

Best Local Similarity

Matches

37.6%; Score 41; DB 1; Length 752;

44.4%; Pred. No. 1.3e+02;

8; Conservative

2; Mismatches

8; Indels

0; Gaps

Query Match


```

Matches      8;  Conservative      2;  Mismatches      6;  Indels      0;  Gaps      0;

QY      1  EGYONYRCRGDDSKV 16
      ||| ||||| :|
Db      233 EGYQNYRGWDQEEKI 248

RESULT 45
B64762
mmp operon transcription regulator mhpR - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C>Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C:Accession: B64762
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: B64762
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-315 <BLAT>
A:Cross-references: UNIPROT:P77569; GB:AE000141; GB:U00096; NID:g1786532; PIDN:AACT73449.
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: mhpR
C:Keywords: transcription regulation

Query Match      36.7%; Score 40; DB 2; Length 315;
Best Local Similarity 50.0%; Pred. No. 82;
Matches      8;  Conservative      2;  Mismatches      6;  Indels      0;  Gaps      0;

QY      1  EGYONYRCRGDDSKV 16
      ||| ||||| :|
Db      233 EGYQNYRGWDQEEKI 248

RESULT 46
A39130
nicotinate phosphoribosyltransferase (EC 2.4.2.11) - Salmonella typhimurium
C:Species: Salmonella typhimurium
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: A39130
R:Vinitsky, A.; Teng, H.; Grubmeyer, C.T.
J. Bacteriol. 173, 536-540, 1991
A:Title: Cloning and nucleic acid sequence of the Salmonella typhimurium pncB gene and s
A:Reference number: A39130; MUID:91100340; PMID:1987148
A:Accession: A39130
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-400 <VIN>
A:Cross-references: UNIPROT:P22253; GB:M55986; NID:g154268; PIDN:AAA27190.1; PID:g154269
C:Genetics:
A:Gene: pncB
C:Function:
A:Pathway: nicotinate and nicotinamide metabolism
C:Superfamily: nicotinate phosphoribosyltransferase
C:Keywords: glycosyltransferase; pentosyltransferase

Query Match      36.7%; Score 40; DB 1; Length 400;
Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches      6;  Conservative      1;  Mismatches      0;  Indels      0;  Gaps      0;

QY      7  YRCRGDD 13
      :|||||
Db      39 FRCRGDD 45

RESULT 47
J00756
nicotinate phosphoribosyltransferase (EC 2.4.2.11) - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

```

```

C:Accession: JQ0756; B64833
R:Wubbolds, M.G.; Terpstra, P.; van Beilen, J.B.; Kingma, J.; Meesters, H.A.R.; Witholt,
J. Biol. Chem. 265, 17665-17672, 1990
A:Title: Variation of cofactor levels in Escherichia coli; sequence analysis and expressi
A:Reference number: JQ0756; MUID:91009224; PMID:2211655
A:Accession: JQ0756
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-400 <WUB>
A:Cross-references: UNIPROT:P18133; GB:J05568; NID:g147306; PIDN:AAA24400.1; PID:g147307
A:Experimental source: strain GEC70
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: B64833
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-400 <BLAT>
A:Cross-references: GB:AE000195; GB:U00096; NID:g1787156; PIDN:AACT4017.1; PID:g1787162;
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: pncB
C:Function:
A:Pathway: nicotinate and nicotinamide metabolism
C:Superfamily: nicotinate phosphoribosyltransferase
C:Keywords: glycosyltransferase; pentosyltransferase

Query Match      36.7%; Score 40; DB 1; Length 400;
Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches      6;  Conservative      1;  Mismatches      0;  Indels      0;  Gaps      0;

QY      7  YRCRGDD 13
      :|||||
Db      39 FRCRGDD 45

RESULT 48
AB0617
nicotinate phosphoribosyltransferase [imported] - Salmonella enterica subsp. enterica ser
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AB0617
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AB0617
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-400 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD05404.1; PID:g16502165; GSPDB:GN00176
C:Genetics:
A:Gene: STY1010
C:Superfamily: nicotinate phosphoribosyltransferase

Query Match      36.7%; Score 40; DB 2; Length 400;
Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches      6;  Conservative      1;  Mismatches      0;  Indels      0;  Gaps      0;

QY      7  YRCRGDD 13
      :|||||
Db      39 FRCRGDD 45

RESULT 49
F90735
nicotinate phosphoribosyltransferase [imported] - Escherichia coli (strain O157:H7, subst

```



```
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: F90755
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gaawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: F90755
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-400 <HAY>
A:Cross-references: UNIPROT:Q8XDE8; GB:BA000007; PIDN:BA834437.1; PID:gl3360473; GSPDB:C
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: ECs1014
C:Superfamily: nicotinate phosphoribosyltransferase

Query Match          36.7%; Score 40; DB 2; Length 400;
Best Local Similarity 85.7%; Pred. NO. 1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      7 YRCRGDD 13
Db      39 FRCRGDD 45

RESULT 50
D85619
nicotinate phosphoribosyltransferase [imported] - Escherichia coli (strain O157:H7, subs
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: D85619
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: D85619
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-400 <STO>
A:Cross-references: UNIPROT:Q8XDE8; GB:AE005174; NID:gl2514106; PIDN:AAG55416.1; GSPDB:C
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: pncB
C:Superfamily: nicotinate phosphoribosyltransferase

Query Match          36.7%; Score 40; DB 2; Length 400;
Best Local Similarity 85.7%; Pred. NO. 1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      7 YRCRGDD 13
Db      39 FRCRGDD 45

Search completed: September 7, 2005, 20:03:45
Job time : 33.3488 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2005, 19:37:54 ; Search time 144.186 Seconds
(without alignments)
53.647 Million cell updates/sec

Title: US-10-812-238B-2

Perfect score: 109

Sequence: 1 EGYIQNYRCRGDDSKVQEAR 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database : A_Geneseq_16Dec04:*

1: Geneseq1980s:*

2: Geneseq1990s:*

3: Geneseq2000s:*

4: Geneseq2001s:*

5: Geneseq2002s:*

6: Geneseq2003as:*

7: Geneseq2003bs:*

8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	109	100.0	311	2 AAW79286	Aaw79286 Human pho
2	109	100.0	311	7 ADE60703	Ade60703 Human pro
3	109	100.0	311	8 ADM72103	Adm72103 Human tau
4	109	100.0	311	8 ADN04415	Adn04415 Antipsori
5	109	100.0	311	8 ADP24274	Adp24274 PRO polyp
6	105	96.3	312	5 ABB99112	Abb99112 Rat phosp
7	105	96.3	312	7 ADE60701	Ade60701 Rat Prote
8	66	60.6	282	5 ABB99111	Abb99111 Rat phosp
9	66	60.6	282	7 ADE58904	Ade58904 Rat Prote
10	66	60.6	282	7 ADE58900	Ade58900 Rat Prote
11	66	60.6	282	7 ADD45532	Add45532 Rat Prote
12	65	59.6	283	7 ADE61256	Ade61256 Rat Prote
13	59	54.1	221	7 ADJ95120	Adj95120 Novel NOV
14	59	54.1	221	7 ADJ95118	Adj95118 Novel NOV
15	59	54.1	284	2 AAW79284	Aaw79284 Human pho
16	59	54.1	285	2 AAW79285	Aaw79285 Human pho
17	59	54.1	285	2 AAY24916	Aay24916 Human pho
18	59	54.1	285	8 ADM72102	Adm72102 Human tau
19	59	54.1	289	7 ADN95432	Adn95432 Human BEC
20	59	54.1	289	8 ADL06556	Adl06556 Human tnm
21	59	54.1	369	8 ADR66929	Adr66929 Human pro
22	59	54.1	369	8 ADR66031	Adr66031 Human pro
23	59	54.1	371	8 ADR66930	Adr66930 Human pro
24	59	54.1	371	8 ADR66032	Adr66032 Human pro
25	49	45.0	847	8 ADS24124	Ads24124 Bacterial

26	49	45.0	1327	8	ADS24139	Ads24139 Bacterial
27	46.5	42.7	705	7	ADB74255	Adb74255 Mycobacte
28	46	42.2	340	4	ABB62735	Abb62735 Drosophi
29	45	41.3	84	4	AAU60236	Aau60236 Propionib
30	45	41.3	84	6	ABW56755	Abw56755 Propionib
31	45	41.3	135	8	ADL92508	Adl92508 Inabill v
32	45	41.3	135	8	ADR41823	Adr41823 versatille
33	45	41.3	822	2	AAW99087	Aaw99087 Human ser
34	45	41.3	875	2	AAW83361	Aaw83361 Human neu
35	45	41.3	875	8	ADQ21226	Adq21226 Human sof
36	45	41.3	875	6	ADQ89082	Adq89082 Human uro
37	44	40.4	191	6	ABU54593	Abu54593 Human NOV
38	44	40.4	385	7	ADE58277	Ade58277 Human pro
39	44	40.4	485	3	ADM04173	Adm04173 Human tra
40	44	40.4	498	3	AAY79209	Aay79209 Human tra
41	44	40.4	498	7	ADN95772	Adn95772 Human BEC
42	44	40.4	498	8	ADP23807	Adp23807 PRO polyp
43	44	40.4	1100	8	ADS24143	Ads24143 Bacterial
44	44	40.4	1907	6	ABG62071	Abg62071 Human ske
45	43	39.4	152	4	ABG23434	Abg23434 Novel hum
46	43	39.4	237	4	AAU28253	Aau28253 Novel hum
47	43	39.4	237	8	ADS12199	Ads12199 Human the
48	43	39.4	371	8	ADN72921	Adn72921 Thale cre
49	43	39.4	421	8	ADJ50947	Adj50947 Human nov
50	43	39.4	424	8	ADJ50945	Adj50945 Human nov
51	43	39.4	428	6	ADA83857	Ada83857 Human POM
52	43	39.4	481	8	ADR09234	Adr09234 Human pro
53	43	39.4	481	8	ADS10966	Ads10966 Human the
54	43	39.4	489	7	ADM04303	Adm04303 Human pro
55	43	39.4	505	8	ADS43566	Ads43566 Bacterial
56	43	39.4	557	4	AAU37211	Aau37211 Staphyloc
57	43	39.4	565	6	ADA89639	Ada89639 Staphyloc
58	43	39.4	565	6	ABU15947	Abu15947 Protein e
59	43	39.4	565	6	ABM71598	Abm71598 Staphyloc
60	43	39.4	566	2	AAY49068	Aay49068 Amino aci
61	43	39.4	566	4	AAE31332	Aae31332 Amino aci
62	43	39.4	566	7	ABW01845	Abw01845 Staphyloc
63	43	39.4	840	4	ABG08798	Abg08798 Novel hum
64	42	38.5	92	4	AAU61299	Aau61299 Propionib
65	42	38.5	92	6	ABM57818	Abm57818 Propionib
66	42	38.5	130	6	ABU50347	Abu50347 Protein e
67	42	38.5	137	7	ADB65039	Adb65039 Human pro
68	42	38.5	157	4	ABE62938	Abbe62938 Drosophi
69	42	38.5	167	7	ABO80468	Abos80468 Pseudomon
70	42	38.5	305	8	ADR41635	Adr41635 Lipid acy
71	42	38.5	305	8	ADR41585	Adr41585 Lipid acy
72	42	38.5	393	8	ADN48160	Adn48160 Thermococ
73	42	38.5	554	6	ABU39968	Abu39968 Protein e
74	42	38.5	1166	5	ABP73801	Abp73801 Candida a
75	42	38.5	1304	6	ABP72189	Abp72189 Plasmodiu
76	42	38.5	3115	4	ABB59485	Abb59485 Drosophi
77	41	37.6	64	4	ABB17072	Abb17072 Human ner
78	41	37.6	129	8	ADK16589	Adk16589 Nanoarcha
79	41	37.6	130	6	ABU40522	Abu40522 Protein e
80	41	37.6	132	7	ADF06937	Adf06937 Bacterial
81	41	37.6	190	5	AAE18631	Aae18631 Equine hi
82	41	37.6	201	2	AAY27061	Aay27061 Recombina
83	41	37.6	236	2	AAV27060	Aav27060 Equine Fc
84	41	37.6	255	2	AAV27058	Aav27058 Equine Fc
85	41	37.6	272	7	ABO67231	Abos67231 Klebsiell
86	41	37.6	290	8	ABO84992	Abos84992 Murine ca
87	41	37.6	396	4	AAE09035	Aae09035 Equine in
88	41	37.6	396	4	AAE09034	Aae09034 Equine in
89	41	37.6	451	4	ABB67864	Abb67864 Drosophi
90	41	37.6	752	5	AAE23119	Aae23119 Influenza
91	41	37.6	757	4	AAE09037	Aae09037 Equine in
92	40.5	37.2	314	7	ABO71590	Abos71590 Pseudomon
93	40.5	37.2	371	4	ABB65431	Abb65431 Drosophi
94	40.5	37.2	373	5	ABP79651	Abp79651 Invertebr
95	40	36.7	20	5	ABP56183	Abp56183 Chimeric
96	40	36.7	94	2	AAV35872	Aav35872 Chlamydia
97	40	36.7	131	6	ABU23041	Abu23041 Protein e
98	40	36.7	132	6	ABU48505	Abu48505 Protein e

99	40	36.7	135	8	ADP29987	Adp29987 Human sec	172	39.5	36.2	211	7	ADB70322	Adb70322 Claudin-7
100	40	36.7	140	4	ABBI1487	Abbi1487 Human sec	173	39.5	36.2	211	7	ADD45212	Add45212 Human Pro
101	40	36.7	246	3	AAAG1453	Aaag1453 Arabidops	174	39.5	36.2	211	8	ADG32539	Adg32539 Human Cla
102	40	36.7	256	5	ABPA0433	Abpa0433 Staphyloc	175	39.5	36.2	211	8	ADJ37101	Adj37101 Human mal
103	40	36.7	256	8	ADSO5965	Adso5965 Staphyloc	176	39.5	36.2	211	8	ADI47185	Adi47185 Human cla
104	40	36.7	301	8	ADS26705	Ads26705 Bacterial	177	39.5	36.2	211	8	ADL26768	Adl26768 Human cla
105	40	36.7	301	8	ADS27091	Ads27091 Bacterial	178	39.5	36.2	212	3	AABS2059	Aabs2059 Human sec
106	40	36.7	302	8	ADS26338	Ads26338 Bacterial	179	39.5	36.2	225	2	AAV60154	Aav60154 Human end
107	40	36.7	387	4	ABGI14658	Abgi14658 Novel hum	180	39.5	36.2	247	5	AAQ75499	Aaq75499 Human col
108	40	36.7	396	4	AAEO9033	Aae09033 Equine in	181	39.5	36.2	247	4	ABP41921	Abp41921 Human ova
109	40	36.7	396	4	AAEO9032	Aae09032 Equine in	182	39.5	36.2	286	6	ABU41431	Abu41431 Protein e
110	40	36.7	400	6	ABU31582	Abu31582 Protein e	183	39.5	36.2	385	8	ADS17418	Ads17418 Amino aci
111	40	36.7	400	6	ABU45262	Abu45262 Protein e	184	39.5	36.2	439	4	ABE70300	Abbe70300 Drosophil
112	40	36.7	400	6	ABU48051	Abu48051 Protein e	185	39.5	36.2	453	8	ADN21519	Adn21519 Bacterial
113	40	36.7	400	6	ABU15369	Abu15369 Protein e	186	39.5	36.2	456	5	ABP38814	Abp38814 Staphyloc
114	40	36.7	401	3	AAAG1452	Aaag1452 Arabidops	187	39.5	36.2	1238	7	ABO77617	Abot77617 Pseudomon
115	40	36.7	418	4	AAU43594	Aau43594 Propionib	188	39	35.8	35	5	ADE01639	Ade01639 Hybrid po
116	40	36.7	418	6	ABM40113	Abm40113 Propionib	189	39	35.8	74	4	ABO40096	Abbo40096 Human mus
117	40	36.7	422	3	AAAG1451	Aaag1451 Arabidops	190	39	35.8	74	6	ABU13390	Abu13390 Novel hum
118	40	36.7	422	5	ABR90818	Abbr90818 Herbicida	191	39	35.8	74	8	ADJ29416	Adj29416 Human mus
119	40	36.7	424	7	ABO65513	Aboc65513 Klebsiell	192	39	35.8	104	4	AAH86759	Aaah86759 Human imm
120	40	36.7	434	6	AAE33569	Aae33569 IL-IRI-CF	193	39	35.8	106	2	AAH60063	Aay60063 Human end
121	40	36.7	440	8	ADN20923	Adn20923 Bacterial	194	39	35.8	122	5	ABP35331	Abp35331 Human ORF
122	40	36.7	454	4	AAAB31968	Aaab31968 Amino aci	195	39	35.8	127	5	ABG70741	Abg70741 Variable
123	40	36.7	480	4	AAAB31962	Aaab31962 Amino aci	196	39	35.8	131	6	ABU22134	Abu22134 Protein e
124	40	36.7	530	5	ABR90929	Abbr90929 Herbicida	197	39	35.8	132	4	AAQ90344	Aaq90344 C Glutami
125	40	36.7	531	4	AAAB31960	Aaab31960 Amino aci	198	39	35.8	132	7	ADJ87359	Adj87359 DNA repli
126	40	36.7	531	8	ADJ45219	Adi45219 Rice isop	199	39	35.8	169	7	ABO70627	Abot70627 Pseudomon
127	40	36.7	536	4	AAAB31964	Aaab31964 Amino aci	200	39	35.8	173	3	AAAB13232	Aabi13232 Caenorhab
128	40	36.7	557	4	ABBE2937	Abbe2937 Drosophil	201	39	35.8	180	4	AAU43388	Aau43388 Propionib
129	40	36.7	710	7	ABO70541	Abot70541 Pseudomon	202	39	35.8	180	6	ABM39907	Abm39907 Propionib
130	40	36.7	730	2	AAAR71379	Aaar71379 Grasshopp	203	39	35.8	185	6	ABU35961	Abu35961 Protein e
131	40	36.7	757	4	AAEO9036	Aae09036 Equine in	204	39	35.8	253	6	ABU36013	Abu36013 Protein e
132	40	36.7	757	5	AAE23109	Aae23109 Influenza	205	39	35.8	288	4	AAU66266	Aau66266 Propionib
133	40	36.7	757	5	ABP53062	Abp53062 Plasmid p	206	39	35.8	288	4	AAU66266	Aau66266 Propionib
134	40	36.7	757	5	ABP53069	Abp53069 Plasmid p	207	39	35.8	297	8	ABU24009	Abu24009 Protein e
135	40	36.7	757	5	ABP53059	Abp53059 FPV-Brati	208	39	35.8	299	8	ADH72218	Adh72218 Human pro
136	40	36.7	757	5	ABP53070	Abp53070 Plasmid p	209	39	35.8	301	4	AAU35607	Aau35607 Haemophil
137	40	36.7	757	5	ABP53067	Abp53067 Plasmid p	210	39	35.8	325	7	ADB65344	Adb65344 Human pro
138	40	36.7	757	5	ABP53060	Abp53060 WSN-PB1 p	211	39	35.8	402	7	ADB65483	Adb65483 Human pro
139	40	36.7	757	5	ABP53064	Abp53064 Plasmid p	212	39	35.8	446	5	AAE18103	Aae18103 Human Cyt
140	40	36.7	757	5	ABP53066	Abp53066 Plasmid p	213	39	35.8	447	5	ABM09643	Abm09643 MutP prot
141	40	36.7	757	5	ABP53068	Abp53068 Plasmid p	214	39	35.8	447	6	ABG72541	Abg72541 Streptoco
142	40	36.7	757	5	ABP53061	Abp53061 Plasmid p	215	39	35.8	461	7	ADG33762	Adg33762 Actinomyc
143	40	36.7	757	5	ABP53065	Abp53065 Plasmid p	216	39	35.8	462	5	ABB47826	Abb47826 Listeria
144	40	36.7	757	5	ABP53063	Abp53063 Plasmid p	217	39	35.8	466	8	ADN20343	Adn20343 Bacterial
145	40	36.7	757	5	ABP53071	Abp53071 Plasmid p	218	39	35.8	470	8	ADN20289	Adn20289 Bacterial
146	40	36.7	757	5	ABBO5765	Abbo5765 Influenza	219	39	35.8	474	6	ABU11478	Abui11478 Human MDD
147	40	36.7	823	6	AAE33571	Aae33571 IL-IRI-CF	220	39	35.8	488	6	ABU11731	Abui11731 Human MDD
148	40	36.7	980	7	ABO82092	Abot82092 Pseudomon	221	39	35.8	509	6	AAU62933	Aau62933 Propionib
149	40	36.7	1235	4	ABG25570	Abg25570 Novel hum	222	39	35.8	509	6	ABM59452	Abm59452 Propionib
150	40	36.7	1299	6	ABP72191	Abp72191 Plasmidii	223	39	35.8	543	5	ABB91383	Abb91383 Herbicida
151	40	36.7	2847	8	ADP25446	Adp25446 Plasmodiu	224	39	35.8	550	7	ADM05649	Adm05649 Human pro
152	40	36.7	5533	4	ABBE5772	Abbe5772 Drosophil	225	39	35.8	557	4	AAU36572	Aau36572 Staphyloc
153	40	36.7	5560	4	ABB71160	Abb71160 Drosophil	226	39	35.8	563	7	AAU33968	Aau33968 Staphyloc
154	40	36.7	5560	8	ADOO1004	Adoo1004 Fruit fly	227	39	35.8	565	7	ADJ769920	Adj769920 Human hea
155	39.5	36.2	84	3	AAAG32835	Aaag32835 Zea mays	228	39	35.8	594	4	ABG06747	Abg06747 Novel hum
156	39.5	36.2	106	3	AAAG43871	Aaag43871 Zea mays	229	39	35.8	602	6	ADA55582	Ada55582 Human pro
157	39.5	36.2	125	3	AAAG32834	Aaag32834 Zea mays	230	39	35.8	614	8	ADR99920	Adr99920 Immune Re
158	39.5	36.2	137	3	AAAY75996	Aay75996 Human ski	231	39	35.8	615	8	ADQ39859	Adq39859 Human myo
159	39.5	36.2	137	4	AAAY55935	Aay55935 Skin cell	232	39	35.8	616	8	ADR99921	Adr99921 Immune Re
160	39.5	36.2	137	5	ABD72135	Abd72135 Human pro	233	39	35.8	652	6	ADB08542	Adb08542 Alloiooc
161	39.5	36.2	191	7	ADP45210	Adp45210 Rat Prote	234	39	35.8	654	8	ADN2472	Adn2472 Bacterial
162	39.5	36.2	194	3	AAAB43554	Aaab43554 Human can	235	39	35.8	659	6	ADB08544	Adb08544 Alloiooc
163	39.5	36.2	205	3	AAAB52101	Aaab52101 Human sec	236	39	35.8	660	2	AAE69633	Aae69633 Human int
164	39.5	36.2	210	2	AAAW63697	Aaw63697 Human sec	237	39	35.8	662	2	AAE69632	Aae69632 Human int
165	39.5	36.2	210	7	ADN95565	Adn95565 Human BEC	238	39	35.8	662	2	AAW12772	Aaw12772 Human int
166	39.5	36.2	211	3	AAAY57609	Aay57609 Human apo	239	39	35.8	662	3	AAAY83950	Aay83950 Human int
167	39.5	36.2	211	5	AAAB69999	Aaab69999 Human Cla	240	39	35.8	662	7	ADE39706	Ade39706 Human int
168	39.5	36.2	211	5	AAO20544	Aao20544 Human Cla	241	39	35.8	662	8	ADK90610	Adk90610 Human IL-
169	39.5	36.2	211	5	ADJ16800	Adj16800 Human NOV	242	39	35.8	662	8	ADM33833	Adm33833 Human IL-
170	39.5	36.2	211	5	ADJ16799	Adj16799 Human NOV	243	39	35.8	662	8	ADQ39861	Adq39861 Human myo
171	39.5	36.2	211	6	ABJ37066	Abj37066 Human bre	244	39	35.8	662	8	ADQ39860	Adq39860 Human myo

245	39	35.8	675	8	ADS11029	Adsl1029 Human the	318	38	34.9	22	4	AAM67400	Aam67400 Human bon
246	39	35.8	690	8	ADS44300	Ads44300 Bacterial	319	38	34.9	22	4	AAM55015	Aam55015 Human bra
247	39	35.8	692	8	ADS11028	Adsl1028 Human the	320	38	34.9	22	4	ABG49059	Abg49059 Human liv
248	39	35.8	694	8	ADO26838	Ado26838 Human rec	321	38	34.9	22	4	AAM02974	Aam02974 Peptide #
249	39	35.8	702	8	ADQ39863	Adq39863 Human myo	322	38	34.9	22	5	ABG37027	Abg37027 Human pep
250	39	35.8	707	6	ABM65866	Abm65866 Propionib	323	38	34.9	35	3	AAY88793	Aay88793 Core poly
251	39	35.8	709	6	ABM08965	Abm08965 Human pro	324	38	34.9	59	4	AAM49621	Aam49621 Propionib
252	39	35.8	748	8	ADQ89640	Adq89640 Antagonis	325	38	34.9	59	6	ABM46140	Abm46140 Propionib
253	39	35.8	754	8	ADQ39862	Adq39862 Human myo	326	38	34.9	77	5	ABP42738	Abp42738 Human ova
254	39	35.8	784	4	ABBG62206	Abbg62206 Drosophil	327	38	34.9	108	7	ADM05267	Adm05267 Human pro
255	39	35.8	790	4	ABBS8707	Abbs8707 Drosophil	328	38	34.9	115	6	ABB70282	Abb70282 Drosophil
256	39	35.8	801	4	AABU34022	Aabu34022 Staphyloc	329	38	34.9	116	6	ABU23431	Abu23431 Protein e
257	39	35.8	801	6	AAMU36926	Aamu36926 Staphyloc	330	38	34.9	120	3	AG02326	Ag02326 Human sec
258	39	35.8	801	6	ABU16301	Abu16301 Protein e	331	38	34.9	129	5	AG93128	Ag93128 S. cerevi
259	39	35.8	801	6	ABM70889	Abm70889 Staphyloc	332	38	34.9	132	4	ABG22111	Abg22111 Novel hum
260	39	35.8	809	6	ABU33713	Abu33713 Protein e	333	38	34.9	136	2	AAM73460	Aam73460 Human sec
261	39	35.8	810	8	ADH72220	Adh72220 Human pro	334	38	34.9	164	5	AAE14398	Aae14398 Human thr
262	39	35.8	898	5	AAU09878	Aau09878 Novel hum	335	38	34.9	172	6	ABU26330	Abu26330 Protein e
263	39	35.8	912	2	NAR75206	Nar75206 Rabbit te	336	38	34.9	172	8	ADQ66177	Adq66177 Novel hum
264	39	35.8	912	7	ADE97397	Ade97397 Rabbit te	337	38	34.9	182	6	ABU94238	Abu94238 Trifolium
265	39	35.8	917	2	AAW00930	Aaw00930 Rat ICAM-	338	38	34.9	206	7	ABM85711	Abm85711 Human pro
266	39	35.8	917	2	AAW60160	Aaw60160 Rat inter	339	38	34.9	213	7	ABO73854	Ab073854 Pseudomon
267	39	35.8	917	2	AAW44836	Aaw44836 Rat ICAM-	340	38	34.9	240	6	ABR43219	Abr43219 Human IRA
268	39	35.8	917	2	AAW59003	Aaw59003 Rat ICAM-	341	38	34.9	242	8	ADK65878	Adk65878 Tobacco N
269	39	35.8	917	2	AAU05465	Aau05465 Rat ICAM-	342	38	34.9	250	6	ABU30346	Abu30346 Protein e
270	39	35.8	917	2	AAW73512	Aaw73512 Rat ICAM-	343	38	34.9	265	8	ADK46930	Adk46930 Streptoco
271	39	35.8	917	7	ADE97409	Ade97409 Norway ra	344	38	34.9	266	4	AAE03975	Aae03975 Human gen
272	39	35.8	917	7	ADE97398	Ade97398 Murine in	345	38	34.9	284	7	ADH85709	Adh85709 Enterococ
273	39	35.8	924	2	AAW00931	Aaw00931 Human ICA	346	38	34.9	285	4	AB73497	Ab73497 Human tra
274	39	35.8	924	2	AAW60161	Aaw60161 Human int	347	38	34.9	303	5	ABP60919	Abp60919 Streptoco
275	39	35.8	924	2	AAW59006	Aaw59006 Human ICA	348	38	34.9	303	6	ABU01907	Abu01907 S. pneumo
276	39	35.8	924	2	AAU05467	Aau05467 Human neu	349	38	34.9	307	8	ABO84993	Ab084993 Human can
277	39	35.8	924	7	AAW73511	Aaw73511 Human sof	350	38	34.9	322	8	ADS21720	Ads21720 Bacterial
278	39	35.8	924	7	ADE97389	Ade97389 Human int	351	38	34.9	326	6	ABU24002	Abu24002 Protein e
279	39	35.8	924	8	ADQ19157	Adq19157 Human sof	352	38	34.9	335	3	AG17233	Ag17233 Arabidops
280	39	35.8	924	8	ADP45605	Adp45605 Human int	353	38	34.9	341	4	ABG22110	Abg22110 Novel hum
281	39	35.8	1004	6	ABU54557	Abu54557 Human NOV	354	38	34.9	347	2	AAW64537	Aaw64537 Human liv
282	39	35.8	1004	8	ADH72222	Adh72222 Human pro	355	38	34.9	347	2	AAM68200	Aam68200 Human sca
283	39	35.8	1004	8	ADH72214	Adh72214 Human pro	356	38	34.9	347	2	AAU13369	Aau13369 Amino aci
284	39	35.8	1004	8	ADH72224	Adh72224 Human pro	357	38	34.9	347	3	ADK78468	Adk78468 Human PRO
285	39	35.8	1137	4	ABB64738	Abb64738 Drosophil	358	38	34.9	347	4	AB880237	Ab880237 Human PRO
286	39	35.8	1137	7	ADC52077	Adc52077 Rice line	359	38	34.9	347	5	AAU83648	Aau83648 Human PRO
287	39	35.8	1269	7	ADD46237	Add46237 Rat Prote	360	38	34.9	347	5	ABB84828	Abb84828 Human PRO
288	39	35.8	1272	7	ADD46239	Add46239 Human Pro	361	38	34.9	347	5	AB95434	Ab95434 Human ang
289	39	35.8	1403	6	ABR52876	AbR52876 Protein s	362	38	34.9	347	6	ABU71615	Abu71615 Human PRO
290	39	35.8	1403	6	ADK62358	Adk62358 Disease t	363	38	34.9	347	6	ABU71470	Abu71470 Human PRO
291	39	35.8	2204	3	AAU51233	Aau51233 Newcastle	364	38	34.9	347	6	ABU80795	Abu80795 Human PRO
292	39	35.8	3069	5	AAE20787	Aae20787 Human C3b	365	38	34.9	347	6	ABO33761	Ab033761 Novel hum
293	39	35.8	3069	5	AAE20900	Aae20900 Human C3b	366	38	34.9	347	6	ABU71316	Abu71316 Human sec
294	39	35.8	3100	5	AAE20789	Aae20789 Human C3b	367	38	34.9	347	6	ABO01799	Ab001799 Novel hum
295	39	35.8	3100	5	AAE20901	Aae20901 Human C3b	368	38	34.9	347	6	ABU54372	Abu54372 Human sec
296	39	35.8	3399	7	ADJ70480	Adj70480 Human hea	369	38	34.9	347	6	ABO47387	Ab047387 Human sec
297	39	35.8	3567	8	ADH72216	Adh72216 Human pro	370	38	34.9	347	6	ABU64524	Abu64524 Human sec
298	38.5	35.3	89	4	ABO55811	Ab055811 Human gen	371	38	34.9	347	6	ABU67370	Abu67370 Human sec
299	38.5	35.3	114	4	AAU29298	Aau29298 C. glutami	372	38	34.9	347	6	ABU89766	Abu89766 Protein d
300	38.5	35.3	149	3	AAU69476	Aau69476 Amino aci	373	38	34.9	347	6	ABO14890	Ab014890 Human sec
301	38.5	35.3	159	7	ADC27626	Adc27626 Human col	374	38	34.9	347	6	ABU82104	Abu82104 Novel hum
302	38.5	35.3	159	7	ADM05361	Adm05361 Human pro	375	38	34.9	347	6	ABU69647	Abu69647 Novel hum
303	38.5	35.3	274	2	ADC99096	Adc99096 Human Kpp	376	38	34.9	347	6	ABO14829	Ab014829 Human sec
304	38.5	35.3	276	2	AAW79287	Aaw79287 Human pro	377	38	34.9	347	6	ABO29353	Ab029353 Human sec
305	38.5	35.3	288	8	ABM61975	Abm61975 Tumour-as	378	38	34.9	347	6	ABJ72284	Abj72284 Human PRO
306	38.5	35.3	725	8	ADG67809	Adg67809 Human ret	379	38	34.9	347	6	ADA18209	Ada18209 Human sec
307	38.5	35.3	897	8	ADI45373	Adi45373 Rice isop	380	38	34.9	347	6	ABO32781	Ab032781 Human sec
308	38.5	35.3	902	4	AAU933269	Aau933269 Human pro	381	38	34.9	347	6	ABO34841	Ab034841 Human PRO
309	38.5	35.3	943	4	AAU94457	Aau94457 Human pro	382	38	34.9	347	6	ADA16184	Ada16184 Human sec
310	38.5	35.3	1275	6	AAU25348	Aau25348 Protein e	383	38	34.9	347	6	ADA42329	Ada42329 Human sec
311	38.5	35.3	1316	4	AAU31884	Aau31884 Novel hum	384	38	34.9	347	6	ABJ72412	Abj72412 Human PRO
312	38.5	35.3	1893	6	AAE33681	Aae33681 Human str	385	38	34.9	347	6	ABO34307	Ab034307 Human sec
313	38	34.9	22	4	AAU15231	Aau15231 Peptide #	386	38	34.9	347	6	ADA16608	Ada16608 Human sec
314	38	34.9	22	4	ABB34221	Abb34221 Peptide #	387	38	34.9	347	7	ADA13037	Ada13037 Human sec
315	38	34.9	22	4	AAW27692	Aaw27692 Peptide #	388	38	34.9	347	7	ADA11905	Ada11905 Human sec
316	38	34.9	22	4	ABB29060	Abb29060 Peptide #	389	38	34.9	347	7	ADA17252	Ada17252 Human sec
317	38	34.9	22	4	ABB19657	Abb19657 Protein #	390	38	34.9	347	7		

391	38	34.9	347	7	ADA42755	Human sec	464	38	34.9	347	8	ADD74323	Human PRO
392	38	34.9	347	7	ABJ72114	Human mem	465	38	34.9	347	8	ADD76053	Novel hum
393	38	34.9	347	7	ABO17580	Human PRO	466	38	34.9	347	8	ADD85545	Novel hum
394	38	34.9	347	7	ADB83604	Novel hum	467	38	34.9	347	8	ADE05094	Human PRO
395	38	34.9	347	7	ADB80710	Novel hum	468	38	34.9	347	8	ADD75307	Human PRO
396	38	34.9	347	7	ADB73251	Novel hum	469	38	34.9	347	8	ADD76851	Novel hum
397	38	34.9	347	7	ADB78333	Novel hum	470	38	34.9	347	8	ADD86619	Novel hum
398	38	34.9	347	7	ADB84981	Human PRO	471	38	34.9	347	8	AD873861	Human sec
399	38	34.9	347	7	ADB77674	Human sec	472	38	34.9	347	8	ADD78087	Novel hum
400	38	34.9	347	7	ADB78087	Novel hum	473	38	34.9	347	8	ADD77595	Novel hum
401	38	34.9	347	7	ADB74810	Human sec	474	38	34.9	347	8	ADD77841	Novel hum
402	38	34.9	347	7	ADB87153	Human PRO	475	38	34.9	347	8	ADD85299	Novel hum
403	38	34.9	347	7	ADB84735	Human PRO	476	38	34.9	347	8	ADD73831	Human PRO
404	38	34.9	347	7	ADB83850	Novel hum	477	38	34.9	347	8	ADD74569	Human PRO
405	38	34.9	347	7	ADB73005	Novel hum	478	38	34.9	347	8	ADD77097	Novel hum
406	38	34.9	347	7	ADC28456	Human sec	479	38	34.9	347	8	ADD85791	Novel hum
407	38	34.9	347	7	ADC39656	Human sec	480	38	34.9	347	8	ADE05340	Human PRO
408	38	34.9	347	7	ADC40170	Human sec	481	38	34.9	347	8	ADD74815	Human PRO
409	38	34.9	347	7	ADC18998	Human sec	482	38	34.9	347	8	ADE99415	Human sec
410	38	34.9	347	7	ADC34294	Human sec	483	38	34.9	347	8	ADE98534	Human sec
411	38	34.9	347	7	ADC29349	Human sec	484	38	34.9	347	8	ADE98961	Human sec
412	38	34.9	347	7	ADC28880	Human sec	485	38	34.9	347	8	ADG05627	Novel hum
413	38	34.9	347	7	ADC40765	Human sec	486	38	34.9	347	8	ADG40431	Human sec
414	38	34.9	347	7	ADC19422	Human sec	487	38	34.9	347	8	ADF73825	Human PRO
415	38	34.9	347	7	ADC33870	Human sec	488	38	34.9	347	8	ADG27181	Human PRO
416	38	34.9	347	7	ADC12940	Human sec	489	38	34.9	347	8	ADF73401	Human sec
417	38	34.9	347	7	ADC36843	Human PRO	490	38	34.9	347	8	ADG11244	Novel hum
418	38	34.9	347	7	ADC21833	Human PRO	491	38	34.9	347	8	ADG12023	Novel hum
419	38	34.9	347	7	ADC49864	Novel hum	492	38	34.9	347	8	ADF94580	Novel hum
420	38	34.9	347	7	ADC49063	Novel hum	493	38	34.9	347	8	ADG06676	Human PRO
421	38	34.9	347	7	ADC49580	Novel hum	494	38	34.9	347	8	ADG92244	Human sec
422	38	34.9	347	7	ADC47441	Novel hum	495	38	34.9	347	8	ADG92671	Human sec
423	38	34.9	347	7	ADC12392	Human sec	496	38	34.9	347	8	ADH39020	Novel hum
424	38	34.9	347	7	ADC47186	Novel hum	497	38	34.9	347	8	ADH20460	Human sec
425	38	34.9	347	7	ADC78061	Novel hum	498	38	34.9	347	8	ADH43457	Human PRO
426	38	34.9	347	7	ADD06296	Novel hum	499	38	34.9	347	8	ADH07315	Human sec
427	38	34.9	347	7	ADD10313	Human sec	500	38	34.9	347	8	ADH59860	Human sec
428	38	34.9	347	7	ADC77815	Novel hum							
429	38	34.9	347	7	ADD04947	Human sec							
430	38	34.9	347	7	ADD11273	Human sec							
431	38	34.9	347	7	ADD03953	Human sec							
432	38	34.9	347	7	ADD03529	Human sec							
433	38	34.9	347	7	ADD50778	Novel hum							
434	38	34.9	347	7	ADD51024	Novel hum							
435	38	34.9	347	7	ADD37066	Human sec							
436	38	34.9	347	7	ADD50505	Human PRO							
437	38	34.9	347	7	ADD50259	Human PRO							
438	38	34.9	347	7	ADD51270	Novel hum							
439	38	34.9	347	7	ADE34781	Human sec							
440	38	34.9	347	7	ADH59264	Human sec							
441	38	34.9	347	7	ADJ38043	Human sec							
442	38	34.9	347	7	ADJ26311	Human sec							
443	38	34.9	347	8	ADC48817	Novel hum							
444	38	34.9	347	8	ADE79226	Human sec							
445	38	34.9	347	8	ADE20988	Novel hum							
446	38	34.9	347	8	ADE05832	Human PRO							
447	38	34.9	347	8	ADD75061	Human PRO							
448	38	34.9	347	8	ADD75807	Novel hum							
449	38	34.9	347	8	ADD85039	Novel hum							
450	38	34.9	347	8	ADD86865	Novel hum							
451	38	34.9	347	8	ADE20742	Novel hum							
452	38	34.9	347	8	ADE39039	Novel hum							
453	38	34.9	347	8	ADE79650	Human sec							
454	38	34.9	347	8	ADE05586	Human PRO							
455	38	34.9	347	8	ADD73571	Human PRO							
456	38	34.9	347	8	ADE73326	Human sec							
457	38	34.9	347	8	ADD78411	Novel hum							
458	38	34.9	347	8	ADE41274	Human sec							
459	38	34.9	347	8	ADE21234	Novel hum							
460	38	34.9	347	8	ADD77349	Novel hum							
461	38	34.9	347	8	ADE20496	Novel hum							
462	38	34.9	347	8	ADD75561	Human PRO							
463	38	34.9	347	8	ADD74077	Human PRO							

ALIGNMENTS

RESULT 1

AAW79286
ID AAW79286 standard; protein; 311 AA.XX
AC AAW79286;XX
DT 15-FEB-1999 (first entry)

Human phosphatidic acid phosphatase beta.

XX
KW Phosphatidic acid phosphatase beta; PAP-beta; human; dephosphorylation;
tumour suppressor; cancer; gene therapy.

OS Homo sapiens.

PN WO9846730-A1.

XX
PD 22-OCT-1998.

PF 16-APR-1998; 98WO-US007928.

XX
PR 17-APR-1997; 97US-00842827.XX
PA (CELL-) CELL THERAPEUTICS INC.XX
PI Leung DW, Tompkins CK;XX
DR WPI: 1998-594568/50.XX
DR N-PSDB; AAV69088.XX
PT New nucleic acid encoding human phosphatidic acid phosphatases - used to

PT regulate levels of lipid cellular mediators and in gene therapy of e.g.
 PT cancer.
 XX Claim 8; Fig 3A-B; 62pp; English.
 CC This is the amino acid sequence of human mature phosphatidic acid
 CC phosphatase-beta (PAP-beta), an enzyme that catalyses the conversion of
 CC phosphatidic acid into diacylglycerol. 3 Variants of human PAP, i.e. PAP-
 CC alpha 1 and 2 (see AA79284-85), PAP-beta (see AA79286) and PAP-gamma
 CC (see AA79287) have been identified. The invention provides PAP
 CC polynucleotides (see AAV69086-89) and polypeptides, a method of preparing
 CC PAP in a transformed host cell, and a method of using PAP to
 CC dephosphorylate a substrate, especially lysophosphatidic acid, ceramide 1
 CC -phosphate or sphingosine 1-phosphate, particularly for production of
 CC diacylglycerol, but also monoacylglycerol, ceramide and sphingosine. PAP
 CC is able to control the balance of lipid mediators of cellular activation
 CC and signal transduction. Sequences that encode PAP are potentially
 CC useful, in gene therapy, for treatment of cancer (PAP may be tumour
 CC suppressors. PAP-alpha is expressed at lower levels in cancer cells than
 CC in normal cells of same tissue type), inflammatory disease and diabetes-
 CC associated obesity
 XX
 SQ Sequence 311 AA;
 Query Match 100.0%; Score 109; DB 2; Length 311;
 Best Local Similarity 100.0%; Pred. No. 1.3e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EGYIYNYRCRGDDSKVQEAR 20
 DB 173 EGYIYNYRCRGDDSKVQEAR 192
 RESULT 2
 ADS60703
 ID ADE60703 standard; protein; 311 AA.
 AC ADE60703;
 XX
 XX 29-JAN-2004 (first entry)
 XX Human Protein NP_003704, SEQ ID NO 6615.
 XX
 XX Human; pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; chronic constriction injury; CCI;
 KW spared nerve injury; SNI; Chung.
 XX
 XX Homo sapiens.
 OS
 XX WO2003016475-A2.
 FN
 XX 27-FEB-2003.
 PD
 XX 14-AUG-2002; 2002WO-US025765.
 PF
 XX 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX
 XX (GEHO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 PA
 XX Woolf C, D'urso D, Befort K, Costigan M;
 XX
 XX WPI; 2003-268312/26.
 DR GENBANK; NP_003704.
 DR
 XX New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 XX
 XX Claim 1; Page; 1017pp; English.
 XX
 XX The invention discloses a composition comprising two or more isolated rat

CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a human protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 311 AA;
 Query Match 100.0%; Score 109; DB 7; Length 311;
 Best Local Similarity 100.0%; Pred. No. 1.3e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EGYIYNYRCRGDDSKVQEAR 20
 DB 173 EGYIYNYRCRGDDSKVQEAR 192
 RESULT 3
 ADM72103
 ID ADM72103 standard; protein; 311 AA.
 XX
 XX ADM72103;
 XX
 XX 03-JUN-2004 (first entry)
 XX Human tau-related polypeptide PPAP2B.
 DE
 XX Tau; PPAP2A; PPAP2B; IHPK1; IHPK3; IHPK2; FLJ20530; DJ434014.5; EZF1;
 KW Loc127424; Loc113179; KIAA0826; neuroprotective; nontropic;
 KW antiparkinsonian; tau-protein kinase; human.
 XX
 XX Homo sapiens.
 OS
 XX WO2004022708-A2.
 FN
 XX 18-MAR-2004.
 PD
 XX 04-SEP-2003; 2003WO-US027590.
 PF
 XX 09-SEP-2002; 2002US-0408877P.
 PR
 XX (BGHM) BRIGHAM & WOMENS HOSPITAL INC.
 PA
 XX Feany MB, Shulman JW;
 XX
 XX WPI; 2004-248456/23.
 DR
 XX New substantially pure Tau-related polypeptides and polynucleotides,
 PT useful for diagnosing and/or treating neurological conditions with
 PT aberrant expression of the Tau-related polypeptide, such as Alzheimer's
 PT disease.
 XX

PS Claim 1; SEQ ID NO 2; 62pp; English.

XX The invention relates to tau-related polypeptide consisting essentially
 CC of an amino acid sequence selected from PPAP2A, PPAP2B, IHPK1, IHPK3,
 CC IHPK2, FLJ20530, D434014.5, EFZ1, LOC127424, LOC131179 and KIAA0826
 CC (ADM72102-ADM72112 respectively). The methods and compositions of the
 CC present invention are useful for the diagnosis and/or treatment of
 CC neurological diseases or conditions associated with aberrant expression
 CC or activity of the Tau-related polypeptide, such as Alzheimer's disease
 CC and Parkinson's disease. The present sequence represents a human tau-
 CC related polypeptide homologue PPAP2B.

XX Sequence 311 AA;

Query Match 100.0%; Score 109; DB 8; Length 311;
 Best Local Similarity 100.0%; Pred. No. 1.3e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGYIYNYRCRGDDSKVQEAR 20
 |||||
 DB 173 EGYIYNYRCRGDDSKVQEAR 192

RESULT 4
 ADN04415
 ID ADN04415 standard; protein; 311 AA.

XX AC ADN04415;

XX DT 01-JUL-2004 (first entry)

XX DE Antipsoriatic protein sequence #401.

XX KW antipsoriatic; gene therapy; psoriasis; diagnosis.

XX OS Homo sapiens.

XX PN WO2004028479-A2.

XX PD 08-APR-2004.

XX PF 25-SEP-2003; 2003WO-US030907.

XX PR 25-SEP-2002; 2002US-0414006P.

XX PA (GETH) GENENTECH INC.

XX PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
 Wu TD;

XX DR WPI; 2004-305105/28.

XX DR N-PSDB; ADN04414.

XX PT New PRO nucleic acid or polypeptide, useful for preparing a
 PT pharmaceutical composition for diagnosing or treating psoriasis in a
 PT mammal.

XX PS Claim 9; SEQ ID NO 809; 3069pp; English.

XX The invention relates to novel polynucleotide and polypeptides for
 CC treating psoriasis or a sequence having at least 80% identity to the
 CC above sequences. The nucleic acid is useful for preparing a composition
 CC for diagnosing or treating psoriasis in a mammal. This sequence
 CC corresponds to one of the polypeptides of the invention.

XX Sequence 311 AA;

Query Match 100.0%; Score 109; DB 8; Length 311;
 Best Local Similarity 100.0%; Pred. No. 1.3e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGYIYNYRCRGDDSKVQEAR 20
 |||||

DB 173 EGYIYNYRCRGDDSKVQEAR 192

RESULT 5
 ADP24274

ID ADP24274 standard; protein; 311 AA.

XX AC ADP24274;

XX DT 18-NOV-2004 (first entry)

XX PRO polypeptide SEQ ID NO:1452.

XX PRO: antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;
 KW osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;
 KW antiaschmatic; hepatotropic; respiratory; gene therapy; immune system.

OS Unidentified.

XX PN WO2004041170-A2.

XX PD 21-MAY-2004.

XX PF 30-OCT-2003; 2003WO-US034312.

XX PR 01-NOV-2002; 2002US-0423394P.

XX PA (GETH) GENENTECH INC.

XX PI Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;
 PI Wu TD;

XX DR WPI; 2004-419628/39.

XX DR N-PSDB; ADP24273.

XX New PRO polypeptides and polynucleotides, useful for treating e.g.
 PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated
 PT renal disease, or demyelinating diseases of the central or peripheral
 PT nervous system.

XX Claim 7; SEQ ID NO 1452; 2940pp; English.

XX The invention relates to a novel isolated nucleic acid and the PRO
 CC polypeptide encoded by it. A protein of the invention has
 CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
 CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
 CC antiaschmatic, hepatotropic, and respiratory activity. A polynucleotide
 CC of the invention may have a use in gene therapy. The PRO polypeptide, its
 CC agonist, antagonist, or antibody that specifically binds to the
 CC polypeptide is useful for treating an immune related disorder such as
 CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
 CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
 CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
 CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
 CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
 CC disease, a demyelinating disease of the central or peripheral nervous
 CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
 CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
 CC disease, infectious or autoimmune chronic active hepatitis, primary
 CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis, CC
 CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
 CC disease, an autoimmune or immune-mediated skin disease, a bullous skin
 CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
 CC disease, asthma, allergic rhinitis, atopic dermatitis, food
 CC hypersensitivity, urticaria, an immunologic disease of the lung,
 CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
 CC pneumonitis, a transplantation associated disease, graft rejection or
 CC graft-versus-host disease. The present sequence represents a PRO protein
 CC of the invention.

XX Sequence 311 AA;

Query Match 100.0%; Score 109; DB 8; Length 311;

Best Local Similarity 100.0%; Pred. No. 1.3e-08; Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGYIQNYRCRGDDSKVQEAR 20
|||||
Db 173 EGYIQNYRCRGDDSKVQEAR 192

RESULT 6
ABB99112
ID ABB99112 standard; protein; 312 AA.
XX
AC ABB99112;
XX
DT 29-OCT-2002 (first entry)
XX
DE Rat phosphatase #2.
XX
KW Phosphatase; prenlyphosphoric acid; prenyl alcohol; rat.
XX
OS Rattus norvegicus.
XX
PN WO200253751-A1.
XX
PD 11-JUL-2002.
XX
PF 20-DEC-2001; 2001WO-JP011223.
XX
PR 28-DEC-2000; 2000JP-00401515.
XX
PR 28-DEC-2000; 2000JP-00401806.
XX
PA (TOYW) TOYOTA CHUO KENKYUSHO KK.
XX
PA (TOYT) TOYOTA JIDOSHA KK.
XX
PI Tokuhiko K, Muramoto N, Yamada Y, Asami O, Hirai M, Ohto C;
PI Obata S, Muramatsu M;
XX
DR WPI; 2002-643303/69.
DR N-PSDB; ABQ78948.
XX
PT Phosphatase polypeptides and encoded polynucleotides with substrate-specificity for prenlyphosphoric acid, applicable in mass-production of all-trans-prenyl alcohols by culturing transformants.
XX
PS Claim 1; Page 67-68; 93pp; Japanese.
XX
CC The invention relates to a novel phosphatase polypeptide, exhibiting substrate-specificity for not less than 15C prenlyphosphoric acid. The polynucleotides and encoded polypeptides are applicable in mass-production of all-trans-prenyl alcohols. The sequence represents a rat phosphatase polypeptide of the invention
XX
SQ Sequence 312 AA;

Query Match 96.3%; Score 105; DB 5; Length 312;
Best Local Similarity 95.0%; Pred. No. 5.4e-08; Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGYIQNYRCRGDDSKVQEAR 20
|||||
Db 174 EGYIQNYRCRGDDSKVQEAR 193

RESULT 7
ADE60701
ID ADE60701 standard; protein; 312 AA.
XX
AC ADE60701;
XX
DT 29-JAN-2004 (first entry)
XX
DE Rat Protein CAA69106, SEQ ID NO 6613.
XX

KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury; chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS Rattus norvegicus.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR WPI; 2003-268312/26.
DR GENEANK; CAA69106.
XX
PT New composition comprising two or more isolated polypeptides, useful for preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene therapy). The sequence presented is a rat protein (shown in Table 2 of the specification) which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 312 AA;

Query Match 96.3%; Score 105; DB 7; Length 312;
Best Local Similarity 95.0%; Pred. No. 5.4e-08; Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EGYIQNYRCRGDDSKVQEAR 20
|||||
Db 174 EGYIQNYRCRGDDSKVQEAR 193

RESULT 8
ABB99111
ID ABB99111 standard; protein; 282 AA.
XX
AC ABB99111;
XX

XX OS Rattus norvegicus.
 XX WO2003016475-A2.
 XX PD 27-FEB-2003.
 XX PF 14-AUG-2002; 2002WO-US025765.
 XX PR 14-AUG-2001; 2001US-0312147P.
 XX PR 01-NOV-2001; 2001US-0345382P.
 XX PR 26-NOV-2001; 2001US-0333347P.
 XX (GEO) GEN HOSPITAL CORP.
 XX (FARB) BAYER AG.
 XX PI Woolf C, D'urso D, Befort K, Costigan M;
 XX WPI; 2003-268312/26.
 XX DR GENBANK; BAA12335.
 XX PT New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 XX PS Claim 1; Page; 1017pp; English.
 XX CC The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 283 AA;
 Query Match 59.6%; Score 65; DB 7; Length 283;
 Best Local Similarity 50.0%; Pred No. 0.095;
 Matches 10; Conservative 7; Mismatches 3; Indels 0; Gaps 0;
 QY 1 EGYIQNYRCRGDSKVKQEAR 20
 Db 145 DGYIEDYICQGNKVKQEAR 164
 RESULT 13
 ID ADJ95120
 ID ADJ95120 standard; protein; 221 AA.
 XX AC ADJ95120;
 XX DT 06-MAY-2004 (first entry)
 XX

DE Novel NOVX protein sequence #174.
 XX antiidiabetic; anorectic; cardiant; hypotensive; antiarteriosclerotic;
 KW anorectic; virucide; antibacterial; fungicide; protozoacide; nootropic;
 KW neuroprotective; antiParkinsonian; anticonvulsant; osteopathic;
 KW antiarthritic; antiinflammatory; dermatological; antiasthmatic;
 KW antileptic; gene therapy; metabolic disorder; diabetes; obesity;
 KW infectious disease; anorexia; cancer; cardiovascular disease;
 KW hypertension; atherosclerosis; neurodegenerative disorder;
 KW Alzheimer's disease; Parkinson's disease; epilepsy; immune disorder;
 KW osteoarthritis; hematopoietic disorder; inflammatory skin disorder;
 KW asthma; dyslipidemia; neurogenesis; cell differentiation;
 KW cell proliferation; hematopoiesis; wound healing; angiogenesis;
 KW chromosome mapping; tissue typing; pharmacogenomic.
 XX Homo sapiens.
 OS WO2003040325-A2.
 XX PD 15-MAY-2003.
 XX PF 05-NOV-2002; 2002WO-US035464.
 XX PR 05-NOV-2001; 2001US-0338626P.
 PR 06-NOV-2001; 2001US-0333072P.
 PR 09-NOV-2001; 2001US-0348283P.
 PR 15-NOV-2001; 2001US-0335610P.
 PR 16-NOV-2001; 2001US-0338543P.
 PR 20-NOV-2001; 2001US-0331630P.
 PR 20-NOV-2001; 2001US-0331641P.
 PR 21-NOV-2001; 2001US-0332152P.
 PR 27-NOV-2001; 2001US-0333461P.
 PR 28-NOV-2001; 2001US-0333912P.
 PR 28-NOV-2001; 2001US-0334027P.
 PR 29-NOV-2001; 2001US-0334300P.
 PR 30-NOV-2001; 2001US-0334421P.
 PR 30-NOV-2001; 2001US-0334526P.
 PR 04-DEC-2001; 2001US-0336576P.
 PR 04-DEC-2001; 2001US-0336664P.
 PR 07-DEC-2001; 2001US-0338314P.
 PR 07-DEC-2001; 2001US-0338390P.
 PR 10-DEC-2001; 2001US-0339006P.
 PR 10-DEC-2001; 2001US-0339008P.
 PR 11-DEC-2001; 2001US-0339286P.
 PR 01-FEB-2002; 2002US-0353280P.
 PR 01-FEB-2002; 2002US-0353288P.
 PR 04-FEB-2002; 2002US-0354392P.
 PR 04-FEB-2002; 2002US-0354393P.
 PR 04-FEB-2002; 2002US-0354409P.
 PR 27-FEB-2002; 2002US-0359944P.
 PR 27-FEB-2002; 2002US-0360148P.
 PR 05-MAR-2002; 2002US-0361790P.
 PR 05-MAR-2002; 2002US-0361833P.
 PR 05-MAR-2002; 2002US-0361925P.
 PR 05-MAR-2002; 2002US-0362300P.
 PR 05-MAR-2002; 2002US-0362625P.
 PR 13-MAR-2002; 2002US-0364000P.
 PR 13-MAR-2002; 2002US-0364181P.
 PR 13-MAR-2002; 2002US-0364182P.
 PR 13-MAR-2002; 2002US-0364197P.
 PR 13-MAR-2002; 2002US-0364227P.
 PR 17-MAY-2002; 2002US-0381621P.
 PR 28-MAY-2002; 2002US-0383675P.
 PR 17-JUL-2002; 2002US-0396703P.
 PR 06-AUG-2002; 2002US-0401552P.
 PR 07-AUG-2002; 2002US-0401594P.
 PR 07-AUG-2002; 2002US-0401787P.
 PR 15-AUG-2002; 2002US-0403619P.
 PR 20-AUG-2002; 2002US-0404821P.
 PR 23-AUG-2002; 2002US-0405368P.
 PR 23-AUG-2002; 2002US-0405402P.
 PR 23-AUG-2002; 2002US-0405496P.
 PR 23-AUG-2002; 2002US-0405631P.

PR 26-AUG-2002; 2002US-0406125P.
 PR 04-NOV-2002; 2002US-00287226.
 XX
 PA (CURA-) CURAGEN CORP.
 XX Agee ML, Alsobrook JP, Berghs C, Boldog FL, Burgess CE, Chant JS;
 PI Chaudhuri A, Dipippo VA, Edinger SR, Eisen A, Ellerman K;
 PI Gangolli EA, Gorman L, Gerlach VL, Ji W, Kekuda R, Khramtsov NV;
 PI Li L, Malyankar UM, Macdougall JR, Mezes PS, Miller CE, Millet I;
 PI Ooi CE, Ort T, Padigaru M, Patturajan M, Rastelli L, Rieger DK;
 PI Rothenberg ME, Shenoy SG, Spaderna SK, Spytek KA, Taupier RJ;
 PI Vernet CAM, Zerhusen BD, Zhong M;
 XX
 DR WPI; 2003-441551/41.
 DR N-PSDB; ADJ95119.
 XX
 PT New isolated NOVX polypeptides and polynucleotides, useful for
 PT preventing, diagnosing or treating NOVX-associated disorders, e.g.
 PT osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
 PT asthma, or infections.
 XX
 PS Claim 1; SEQ ID NO 348; 800pp; English.
 XX
 CC The invention relates to novel isolated polypeptides, mature forms of
 CC these, or a sequence that is at least 95 % identical to, or having one or
 CC more conservative amino acid substitutions in the polypeptides. The
 CC polypeptides, nucleic acid molecules and antibodies are useful in the
 CC manufacture of a medicament for treating a syndrome associated with a
 CC human disease, preferably a NOVX-associated disorder. The nucleic acid
 CC molecules, polypeptides and antibodies are useful for treating,
 CC preventing or diagnosing diseases such as metabolic disorders, diabetes,
 CC obesity, infectious diseases (viral, bacterial, fungal, helminthic, and
 CC protozoal), anorexia, cancer, cardiovascular diseases (hypertension,
 CC atherosclerosis), neurodegenerative disorders, Alzheimer's disease,
 CC Parkinson's disease, epilepsy, immune disorders (osteoarthritis),
 CC hematopoietic disorders, inflammatory skin disorders, asthma, and various
 CC dyslipidemias. The nucleic acids and polypeptides may also be used as
 CC targets for the identification of small molecules that modulate or
 CC inhibit e.g. neurogenesis, cell differentiation, cell proliferation,
 CC hematopoiesis, wound healing and angiogenesis, in gene therapy, in
 CC generation of antibodies that bind immunospecifically to NOVX substances
 CC for use in therapeutic or diagnostic methods. The nucleic acids are
 CC further used as hybridization probes, in chromosome mapping, tissue
 CC typing, preventive medicine, and pharmacogenomics. This sequence
 CC corresponds to one of the NOVX polypeptides of the invention.
 XX
 SQ Sequence 221 AA;
 Query Match 54.1%; Score 59; DB 7; Length 221;
 Best Local Similarity 50.0%; Pred. No. 0.64;
 Matches 10; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
 QY 1 EGYIONVRCRGDDSKVQEAR 20
 Db 82 DGYIEYICRGNARVKEGR 101
 RESULT 14
 ADJ95118
 ID ADJ95118 standard; protein; 221 AA.
 XX
 AC ADJ95118;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Novel NOVX protein sequence #173.
 XX
 KW antidiabetic; anorectic; cardiant; hypotensive; antiarteriosclerotic;
 KW anorectic; virucide; antibacterial; fungicide; protozoacide; nootropic;
 KW neuroprotective; antiparkinsonian; anticonvulsant; osteopathic;
 KW antiarthritic; antiinflammatory; dermatological; antiasthmatic;
 KW antilipemic; gene therapy; metabolic disorder; diabetes; obesity;
 KW infectious disease; anorexia; cancer; cardiovascular disease;

KW hypertension; atherosclerosis; neurodegenerative disorder;
 KW Alzheimer's disease; Parkinson's disease; epilepsy; immune disorder;
 KW osteoarthritis; hematopoietic disorder; inflammatory skin disorder;
 KW asthma; dyslipidemia; neurogenesis; cell differentiation;
 KW cell proliferation; hematopoiesis; wound healing; angiogenesis;
 KW chromosome mapping; tissue typing; pharmacogenomic.
 XX Homo sapiens.
 OS
 PN WO2003040325-A2.
 XX
 PD 15-MAY-2003.
 XX
 XX 05-NOV-2002; 2002WO-US035464.
 PF
 XX
 XX 05-NOV-2001; 2001US-0338626P.
 PR
 PR 06-NOV-2001; 2001US-033072P.
 PR
 PR 09-NOV-2001; 2001US-0348283P.
 PR
 PR 15-NOV-2001; 2001US-0335610P.
 PR
 PR 16-NOV-2001; 2001US-0338543P.
 PR
 PR 20-NOV-2001; 2001US-0331630P.
 PR
 PR 20-NOV-2001; 2001US-0331641P.
 PR
 PR 21-NOV-2001; 2001US-0332152P.
 PR
 PR 27-NOV-2001; 2001US-0333461P.
 PR
 PR 28-NOV-2001; 2001US-0333912P.
 PR
 PR 28-NOV-2001; 2001US-0334027P.
 PR
 PR 29-NOV-2001; 2001US-0334300P.
 PR
 PR 30-NOV-2001; 2001US-0334421P.
 PR
 PR 30-NOV-2001; 2001US-0334526P.
 PR
 PR 04-DEC-2001; 2001US-0335576P.
 PR
 PR 04-DEC-2001; 2001US-0336664P.
 PR
 PR 07-DEC-2001; 2001US-0338314P.
 PR
 PR 10-DEC-2001; 2001US-0338390P.
 PR
 PR 10-DEC-2001; 2001US-0339006P.
 PR
 PR 11-DEC-2001; 2001US-0339008P.
 PR
 PR 11-DEC-2001; 2001US-0339286P.
 PR
 PR 01-FEB-2002; 2002US-0353280P.
 PR
 PR 01-FEB-2002; 2002US-0353288P.
 PR
 PR 04-FEB-2002; 2002US-0354392P.
 PR
 PR 04-FEB-2002; 2002US-0354393P.
 PR
 PR 04-FEB-2002; 2002US-0354409P.
 PR
 PR 27-FEB-2002; 2002US-0359444P.
 PR
 PR 27-FEB-2002; 2002US-0360148P.
 PR
 PR 05-MAR-2002; 2002US-0361790P.
 PR
 PR 05-MAR-2002; 2002US-0361833P.
 PR
 PR 05-MAR-2002; 2002US-0361925P.
 PR
 PR 05-MAR-2002; 2002US-0362300P.
 PR
 PR 05-MAR-2002; 2002US-0362625P.
 PR
 PR 13-MAR-2002; 2002US-0364000P.
 PR
 PR 13-MAR-2002; 2002US-0364181P.
 PR
 PR 13-MAR-2002; 2002US-0364182P.
 PR
 PR 13-MAR-2002; 2002US-0364197P.
 PR
 PR 13-MAR-2002; 2002US-0364227P.
 PR
 PR 17-MAY-2002; 2002US-0381621P.
 PR
 PR 28-MAY-2002; 2002US-0383675P.
 PR
 PR 17-JUL-2002; 2002US-0396703P.
 PR
 PR 06-AUG-2002; 2002US-0401552P.
 PR
 PR 07-AUG-2002; 2002US-0401594P.
 PR
 PR 07-AUG-2002; 2002US-0401787P.
 PR
 PR 15-AUG-2002; 2002US-0403619P.
 PR
 PR 20-AUG-2002; 2002US-0404821P.
 PR
 PR 23-AUG-2002; 2002US-0405368P.
 PR
 PR 23-AUG-2002; 2002US-0405402P.
 PR
 PR 23-AUG-2002; 2002US-0405496P.
 PR
 PR 23-AUG-2002; 2002US-0405631P.
 PR
 PR 26-AUG-2002; 2002US-0406125P.
 PR
 PR 04-NOV-2002; 2002US-00287226.
 XX
 XX (CURA-) CURAGEN CORP.
 XX Agee ML, Alsobrook JP, Berghs C, Boldog FL, Burgess CE, Chant JS;
 XX Chaudhuri A, Dipippo VA, Edinger SR, Eisen A, Ellerman K;
 PI Gangolli EA, Gorman L, Gerlach VL, Ji W, Kekuda R, Khramtsov NV;
 PI Chaudhuri A, Dipippo VA, Edinger SR, Eisen A, Ellerman K;
 PI Gangolli EA, Gorman L, Gerlach VL, Ji W, Kekuda R, Khramtsov NV;

PI Li L, Malvankar UM, Macdougall JR, Mexes PS, Miller CE, Millet I;
 PI Ooi CE, Ort T, Padigar M, Patturajan M, Rastelli L, Rieger DK;
 PI Rothenberg ME, Shenoy SG, Spaderna SK, Spytek KA, Taupier RJ;
 PI Vernet CAM, Zerhusen BD, Zhong M;
 XX
 DR WPI; 2003-441551/41.
 DR N-PSDB; ADJ95117.
 XX
 XX New isolated NOVX polypeptides and polynucleotides, useful for
 PT preventing, diagnosing or treating NOVX-associated disorders, e.g.
 PT osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
 PT asthma, or infections.
 XX
 PS Claim 1; SEQ ID NO 346; 800pp; English.
 XX
 CC The invention relates to novel isolated polypeptides, mature forms of
 CC these, or a sequence that is at least 95 % identical to, or having one or
 CC more conservative amino acid substitutions in the polypeptides. The
 CC polypeptides, nucleic acid molecules and antibodies are useful in the
 CC manufacture of a medicament for treating a syndrome associated with a
 CC human disease, preferably a NOVX-associated disorder. The nucleic acid
 CC molecules, polypeptides and antibodies are useful for treating,
 CC preventing or diagnosing diseases such metabolic disorders, diabetes,
 CC obesity, infectious diseases (viral, bacterial, fungal, helminthic, and
 CC protozoal), anorexia, cancer, cardiovascular diseases (hypertension,
 CC atherosclerosis), neurodegenerative disorders, Alzheimer's disease,
 CC Parkinson's disease, epilepsy, immune disorders (osteoarthritis),
 CC hematopoietic disorders, inflammatory skin disorders, asthma, and various
 CC dyslipidemias. The nucleic acids and polypeptides may also be used as
 CC targets for the identification of small molecules that modulate or
 CC inhibit e.g. neurogenesis, cell differentiation, cell proliferation,
 CC hematopoiesis, wound healing and angiogenesis, in gene therapy, in
 CC generation of antibodies that bind immunospecifically to NOVX substances
 CC for use in therapeutic or diagnostic methods. The nucleic acids are
 CC further used as hybridization probes, in chromosome mapping, tissue
 CC typing, preventive medicine, and pharmacogenomics. This sequence
 CC corresponds to one of the NOVX polypeptides of the invention.
 XX
 SQ Sequence 221 AA;
 Query Match 54.1%; Score 59; DB 7; Length 221;
 Best Local Similarity 50.0%; Pred. No. 0.64;
 Matches 10; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
 OY 1 EGYIONRCRGDSDSKVQEAR 20
 :|||: |||: :||: |||:
 Db 82 DGYIEYICRGNAERVKGR 101
 RESULT 15
 AAW79284
 ID AAW79284 standard; protein; 284 AA.
 XX
 AC AAW79284;
 XX
 DT 15-FEB-1999 (first entry)
 XX
 DE Human phosphatidic acid phosphatase alpha 1.
 XX
 KW Phosphatidic acid phosphatase alpha 1; PAP-alpha 1; human;
 KW dephosphorylation; tumour suppressor; cancer; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO9846730-A1.
 XX
 PD 22-OCT-1998.
 XX
 PF 16-APR-1998; 98WO-US007928.
 XX
 PR 17-APR-1997; 97US-00842827.
 XX
 PA (CELL-) CELL THERAPEUTICS INC.
 PI Leung DW, Tompkins CK;
 DR WPI; 1998-594568/50.
 DR N-PSDB; AAV69087.

XX
 PI Leung DW, Tompkins CK;
 XX
 DR WPI; 1998-594568/50.
 DR N-PSDB; AAV69086.
 XX
 XX New nucleic acid encoding human phosphatidic acid phosphatases - used to
 PT regulate levels of lipid cellular mediators and in gene therapy of e.g.
 PT cancer.
 XX
 PS Claim 2; Fig 1A-B; 62pp; English.
 XX
 CC This is the amino acid sequence of human mature phosphatidic acid
 CC phosphatase-alpha 1 (PAP-alpha-1), an enzyme that catalyses the
 CC conversion of phosphatidic acid into diacylglycerol. 3 Variants of human
 CC PAP, i.e. PAP-alpha 1 and its alternatively spliced form PAP-alpha 2 (see
 CC AAW79285), PAP-beta (see AAW79286) and PAP-gamma (see AAW79287) have been
 CC identified. The invention provides PAP polynucleotides (see AAV69086-89)
 CC and polypeptides, a method of preparing a PAP polypeptide in a
 CC transformed host cell, and a method of using PAP to dephosphorylate a
 CC substrate, especially lysophosphatidic acid, ceramide 1-phosphate or
 CC sphingosine 1-phosphate, particularly for production of diacylglycerol,
 CC but also monacylglycerol, ceramide and sphingosine. PAP is able to
 CC control the balance of lipid mediators of cellular activation and signal
 CC transduction. Sequences that encode PAP are potentially useful, in gene
 CC therapy, for treatment of cancer (PAP may be tumour suppressors, PAP-
 CC alpha is expressed at lower levels in cancer cells than in normal cells
 CC of same tissue type), inflammatory disease and diabetes-associated
 CC obesity
 XX
 SQ Sequence 284 AA;
 Query Match 54.1%; Score 59; DB 2; Length 284;
 Best Local Similarity 50.0%; Pred. No. 0.84;
 Matches 10; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
 OY 1 EGYIONRCRGDSDSKVQEAR 20
 :|||: |||: :||: |||:
 Db 145 DGYIEYICRGNAERVKGR 164
 RESULT 16
 AAW79285
 ID AAW79285 standard; protein; 285 AA.
 XX
 AC AAW79285;
 XX
 DT 15-FEB-1999 (first entry)
 XX
 DE Human phosphatidic acid phosphatase alpha 2.
 XX
 KW Phosphatidic acid phosphatase alpha 2; PAP-alpha 2; human;
 KW dephosphorylation; tumour suppressor; cancer; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO9846730-A1.
 XX
 PD 22-OCT-1998.
 XX
 PF 16-APR-1998; 98WO-US007928.
 XX
 PR 17-APR-1997; 97US-00842827.
 XX
 PA (CELL-) CELL THERAPEUTICS INC.
 PI Leung DW, Tompkins CK;
 DR WPI; 1998-594568/50.
 DR N-PSDB; AAV69087.
 XX
 XX New nucleic acid encoding human phosphatidic acid phosphatases - used to
 PT regulate levels of lipid cellular mediators and in gene therapy of e.g.

Db 146 DGYIEYICRGNAERVKEGR 165

RESULT 19
ADN95452
ID ADN95452 standard; protein; 289 AA.
XX
AC ADN95452;
XX
DT 01-JUL-2004 (first entry)
XX
DE Human BEC/LEC-related protein sequence SeqID375.
XX
KW growth; differentiation; blood endothelial cell; BEC;
KW lymphatic endothelial cell; LEC; hereditary lymphoedema; VEGFR-3;
KW lymphatic growth agent; VEGF-C; VEGF-D; antiangiogenic; cytostatic;
KW vasotropic; antiinflammatory; gene therapy; endothelial cell disorder;
KW inflammatory disease; cancer metastasis; lymphatic system; human.
XX
OS Homo sapiens.
XX
PN WO2003080640-A1.
XX
PD 02-OCT-2003.
XX
PF 07-MAR-2003; 2003WO-US006900.
XX
PR 07-MAR-2002; 2002US-0363019P.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
PA (LICN) LICENTIA LTD.
XX
PI Alitalo K, Makinen T, Petrova T, Saharinen P, Saharinen J;
XX
DR WPI; 2003-876899/81.
DR N-PSDB; ADN95453.
XX
PS Example 1; SEQ ID NO 375; 176pp; English.

This invention relates to a method of differentially modulating the growth or differentiation of blood endothelial cells (BEC) or lymphatic endothelial cells (LEC) comprises contacting endothelial cells with a composition comprising an agent that differentially modulates blood or lymphatic endothelial cells. Treating hereditary lymphoedema comprises identifying a human subject with lymphoedema and with a mutation in at least one allele of a gene encoding a LEC protein, where the mutation correlates with lymphoedema in human subjects, and with the proviso that the LEC protein is not VEGFR-3; and administering to the subject a composition comprising a lymphatic growth agent selected from VEGF-C or VEGF-D polypeptides and polynucleotides. The invention may be useful for the development of compounds with an antiangiogenic, cytostatic, vasotropic or antiinflammatory activity or for gene therapy. The method is useful in modulating the growth or differentiation of blood endothelial cells or lymphatic endothelial cells, in treating hereditary lymphoedema, in screening for an endothelial cell disorder or predisposition to the disorder or in monitoring the efficacy or toxicity of a drug on endothelial cells. The agent is useful in manufacturing a medicament for the differential modulation of blood vessel endothelial cell or lymphatic vessel endothelial cell growth or differentiation. The lymphatic growth agent may also be used in manufacturing a medicament for the treatment of hereditary lymphoedema resulting from a mutation in a LEC gene or of other diseases involving the lymphatic vessels, such as various inflammatory diseases and cancer metastasis via the lymphatic system. The present sequence is that of a human LEC/BEC differentially expressed protein which is related to the method of the invention. Note: This sequence does not appear in the specification but was obtained by the indexer using the source data given in table 14 of the specification.

XX
SQ Sequence 289 AA;

Query Match 54.1%; Score 59; DB 7; Length 289;
Best Local Similarity 50.0%; Pred. No. 0.86;
Matches 10; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 EGYIONVRCRGDDSKVQEAR 20
:||||:||||:|:|:
Db 150 DGYIEYICRGNAERVKEGR 169

RESULT 20
ADL06556
ID ADL06556 standard; protein; 289 AA.
XX
AC ADL06556;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human tumour-associated antigenic target (TAT) polypeptide #55.
XX
KW Human; tumour-associated antigenic target; TAT; cell death; tumour;
KW cancer; cytostatic.
XX
OS Homo sapiens.
XX
PN WO2004016225-A2.
XX
PD 26-FEB-2004.
XX
PF 19-AUG-2003; 2003WO-US025892.
XX
PR 19-AUG-2002; 2002US-0404809P.
PR 21-AUG-2002; 2002US-0405645P.
PR 23-SEP-2002; 2002US-0413192P.
PR 15-OCT-2002; 2002US-0419008P.
PR 15-NOV-2002; 2002US-0426847P.
PR 02-JUL-2003; 2003US-0484959P.
XX
PA (GETH) GENENTECH INC.
XX
PI Desauvage FJ, Frantz G, Hillan KJ, Polakis P, Polson A, Smith V;
PI Spencer SD, Wu TD, Zhang Z;
XX
DR WPI; 2004-257144/24.
DR N-PSDB; ADL06477.
XX
PT New antibody that binds to a tumor-associated antigenic target (TAT) polypeptide, useful for preparing a composition for diagnosing or treating cancer.
XX
PS Claim 2; SEQ ID NO 136; 319pp; English.

The present invention relates to the isolation of human tumour-associated antigenic target (TAT) polynucleotide and polypeptide sequences. Also disclosed is an antibody that binds to a TAT polypeptide. The antibody is a monoclonal antibody, an antibody fragment, a chimeric antibody or a humanised antibody. It is conjugated to a growth inhibitory agent. It is produced in bacteria or in CHO cells and induces death of a cell to which it binds. The antibody is useful for preparing a composition for diagnosing or treating tumours and cancer. The present sequence represents a human TAT polypeptide of the invention.

XX
SQ Sequence 289 AA;

Query Match 54.1%; Score 59; DB 8; Length 289;
Best Local Similarity 50.0%; Pred. No. 0.86;
Matches 10; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 EGYIONVRCRGDDSKVQEAR 20
:||||:||||:|:|:
Db 150 DGYIEYICRGNAERVKEGR 169

RESULT 21
ADR66929
ID ADR66929 standard; protein; 369 AA.
XX

Matches	10;	Conservative	5;	Mismatches	5;	Indels	0;	Gaps	0;
QY	1	EGYQNYRCRGDDSKVQEAR	20						
		: : : :							
DB	230	DGYIEYYICRGNAERVKGR	249						
RESULT 22									
ADR66031									
ID	ADR66031	standard; protein; 369 AA.							
XX	AC	ADR66031;							
XX	DT	02-DEC-2004 (first entry)							
XX	XX	Human prostatic carcinoma derived protein SEQ ID 227 #1.							
XX	XX	human; cytostatic; diagnosis; prostatic cancer;							
XX	KW	differential expression analysis.							
XX	XX	Homo sapiens.							
XX	OS	WO2004076614-A2.							
XX	PN	10-SEP-2004.							
XX	PP	22-FEB-2004; 2004WO-DE000433.							
XX	XX	27-FEB-2003; 2003DE-01009985.							
XX	PR	14-MAY-2003; 2003DE-01022134.							
XX	XX	(HINZ/) HINZMANN B.							
XX	PA	(DAHL/) DAHL E.							
XX	PA	(ROSE/) ROSENTHAL A.							
XX	PA	(HERM/) HERMANN K.							
XX	PA	(PILA/) PILARSKY C.							
XX	PI	Hinzmann B, Dahl E, Rosenthal A, Hermann K, Pilarsky C, Specht T;							
XX	PI	Schmitt A, Beckmann G, Bruemendorf T, Kinnemann H, Roepcke S;							
XX	PI	Xinzhong L, Staub E;							
XX	XX	WPI; 2004-653386/63.							
XX	DR	New nucleic acids, and encoded proteins, from prostatic cancer tissue,							
XX	PT	useful for diagnosis, treatment and in screening for specific binding							
XX	PT	agents.							
XX	XX	Claim 2; Page 622; 1607pp; German.							
XX	CC	This invention describes novel cytostatic polynucleotide and polypeptide							
XX	CC	sequences which can be used in a method for diagnosing prostatic cancer							
XX	CC	or the risk of developing prostatic cancer. Diagnosis is based on							
XX	CC	determining over transcription or over expression of the sequences in							
XX	CC	prostatic tissue. Screening for inhibitors of the sequences or detection							
XX	CC	substances involves a binding assay, any compounds that bind are							
XX	CC	selected, optionally after deconvolution of mixtures. Detection of a							
XX	CC	predetermined minimum level of the reporter indicates the presence of							
XX	CC	tumour cells. Inhibitors can be chosen from antisense oligonucleotides,							
XX	CC	short-interfering RNA or ribozymes; an organic molecule of molecular							
XX	CC	weight below 5000, preferably 300, that binds to the polypeptide; an							
XX	CC	aptamer against the polypeptide; a (monoclonal) antibody (Ab) against the							
XX	CC	polypeptide, preferably humanised or human; an anti-idiotypic, non-human							
XX	CC	(monoclonal) antibody directed against Ab or any of the above derivatised							
XX	CC	with a reporter group, cell toxin, immunostimulatory molecules and/or							
XX	CC	radioisotope. The polynucleotides are identified in human prostatic							
XX	CC	cancer by differential expression analysis, using DNA microarrays,							
XX	CC	between normal and tumorous tissues, with (over)expression being detected							
XX	CC	by quantitative PCR. Analysis of prostatic cancer samples showed that							
XX	CC	CD24 was upregulated in many of them. Sections of tissue, isolated from							
XX	CC	prostatic cancer patients, or subjects at risk, were incubated							
XX	CC	sequentially with anti-human CD4 murine monoclonal antibodies;							
XX	CC	biotinylated second antibody; streptavidin-conjugated horseradish							
XX	CC	peroxidase and then diaminobenzidine as colour former (brown). The							

CC samples were counterstained with hemalum (blue). Malignant cells stained
 CC strongly but non-malignant cells only weakly. In 15 of 63 samples of
 CC adenocarcinoma, membrane and cytoplasmic staining was very strong, and
 CC lymph node metastases were also stained. ADR65805-ADR66954 represent the
 CC polynucleotide and polypeptide sequences used in the method of the
 CC invention.
 XX
 SQ Sequence 369 AA;
 Query Match 54.1%; Score 59; DB 8; Length 369;
 Best Local Similarity 50.0%; Pred. No. 1.1;
 Matches 10; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
 QY 1 EGYIQNYRCRGDDSKVQEAR 20
 :|||: |||: :||: |
 Db 230 DGYIEYYICRGNAERVKEGR 249
 :|||: |||: :||: |
 RESULT 23
 ADR66930
 ID ADR66930 standard; protein; 371 AA.
 XX
 AC ADR66930;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Human prostatic carcinoma derived DNA SEQ ID 228 #4.
 XX
 KW human; cytostatic; diagnosis; prostatic cancer;
 KW differential expression analysis.
 XX
 OS Homo sapiens.
 XX
 PN WO2004076614-A2.
 XX
 PD 10-SEP-2004.
 XX
 PF 22-FEB-2004; 2004WO-DE000433.
 XX
 PR 27-FEB-2003; 2003DE-01009985.
 PR 14-MAY-2003; 2003DE-01022134.
 XX
 PA (HINZ/) HINZMANN B.
 PA (DAHL/) DAHL E.
 PA (ROSE/) ROSENTHAL A.
 PA (HERM/) HERMANN K.
 PA (PILA/) PILARSKY C.
 XX
 PI Hinzmann B, Dahl E, Rosenthal A, Hermann K, Pilarsky C, Specht T;
 PI Schmitt A, Beckmann G, Bruemendorf T, Kinnemann H, Roepcke S;
 PI Xinzhang L, Staub E;
 XX
 DR WPI; 2004-653386/63.
 XX
 PT New nucleic acids, and encoded proteins, from prostatic cancer tissue,
 PT useful for diagnosis, treatment and in screening for specific binding
 agents.
 XX
 PS Claim 2; Page 1583; 1607pp; German.
 XX
 CC This invention describes novel cytostatic polynucleotide and polypeptide
 CC sequences which can be used in a method for diagnosing prostatic cancer
 CC or the risk of developing prostatic cancer. Diagnosis is based on
 CC determining over transcription or over expression of the sequences in
 CC prostatic tissue. Screening for inhibitors of the sequences or detection
 CC substances involves a binding assay, any compounds that bind are
 CC selected, optionally after deconvolution of mixtures. Detection of a
 CC predetermined minimum level of the reporter indicates the presence of
 CC tumour cells. Inhibitors can be chosen from antisense oligonucleotides,
 CC short-interfering RNA or ribozymes; an organic molecule of molecular
 CC weight below 5000, preferably 300, that binds to the polypeptide; an
 CC aptamer against the polypeptide; a (monoclonal) antibody (Ab) against the
 CC polypeptide, preferably humanised or human; an anti-idiotype, non-human

CC (monoclonal) antibody directed against Ab or any of the above derivatised
 CC with a reporter group, cell toxin, immunostimulatory molecules and/or
 CC radioisotope. The polynucleotides are identified in human prostatic
 CC cancer by differential expression analysis, using DNA microarrays,
 CC between normal and tumorous tissues, with (over)expression being detected
 CC by quantitative PCR. Analysis of prostatic cancer samples showed that
 CC CD24 was upregulated in many of them. Sections of tissue, isolated from
 CC prostatic cancer patients, or subjects at risk, were incubated
 CC sequentially with anti-human CD4 murine monoclonal antibodies;
 CC biotinylated second antibody; streptavidin-conjugated horseradish
 CC peroxidase and then diaminobenzidine as colour former (brown). The
 CC samples were counterstained with hemalum (blue). Malignant cells stained
 CC strongly but non-malignant cells only weakly. In 15 of 63 samples of
 CC adenocarcinoma, membrane and cytoplasmic staining was very strong, and
 CC lymph node metastases were also stained. ADR65805-ADR66954 represent the
 CC polynucleotide and polypeptide sequences used in the method of the
 CC invention.
 XX
 SQ Sequence 371 AA;
 Query Match 54.1%; Score 59; DB 8; Length 371;
 Best Local Similarity 50.0%; Pred. No. 1.1;
 Matches 10; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
 QY 1 EGYIQNYRCRGDDSKVQEAR 20
 :|||: |||: :||: |
 Db 232 DGYIEYYICRGNAERVKEGR 251
 :|||: |||: :||: |
 RESULT 24
 ADR66032
 ID ADR66032 standard; protein; 371 AA.
 XX
 AC ADR66032;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Human prostatic carcinoma derived protein SEQ ID 228 #1.
 XX
 KW human; cytostatic; diagnosis; prostatic cancer;
 KW differential expression analysis.
 XX
 OS Homo sapiens.
 XX
 PN WO2004076614-A2.
 XX
 PD 10-SEP-2004.
 XX
 PF 22-FEB-2004; 2004WO-DE000433.
 XX
 PR 27-FEB-2003; 2003DE-01009985.
 PR 14-MAY-2003; 2003DE-01022134.
 XX
 PA (HINZ/) HINZMANN B.
 PA (DAHL/) DAHL E.
 PA (ROSE/) ROSENTHAL A.
 PA (HERM/) HERMANN K.
 PA (PILA/) PILARSKY C.
 XX
 PI Hinzmann B, Dahl E, Rosenthal A, Hermann K, Pilarsky C, Specht T;
 PI Schmitt A, Beckmann G, Bruemendorf T, Kinnemann H, Roepcke S;
 PI Xinzhang L, Staub E;
 XX
 DR WPI; 2004-653386/63.
 XX
 PT New nucleic acids, and encoded proteins, from prostatic cancer tissue,
 PT useful for diagnosis, treatment and in screening for specific binding
 agents.
 XX
 PS Claim 2; Page 623; 1607pp; German.
 XX
 CC This invention describes novel cytostatic polynucleotide and polypeptide
 CC sequences which can be used in a method for diagnosing prostatic cancer

or the risk of developing prostatic cancer. Diagnosis is based on determining over transcription or over expression of the sequences in prostatic tissue. Screening for inhibitors of the sequences or detection substances involves a binding assay, any compounds that bind are selected, optionally after deconvolution of mixtures. Detection of a predetermined minimum level of the reporter indicates the presence of tumour cells. Inhibitors can be chosen from antisense oligonucleotides, short-interfering RNA or ribozymes; an organic molecule of molecular weight below 5000, preferably 300, that binds to the polypeptide; an aptamer against the polypeptide; a (monoclonal) antibody (Ab) against the polypeptide, preferably humanised or human; an anti-idiotypic, non-human (monoclonal) antibody directed against Ab or any of the above derivatised with a reporter group, cell toxin, immunostimulatory molecules and/or radioisotope. The polynucleotides are identified in human prostatic cancer by differential expression analysis, using DNA microarrays, between normal and tumorous tissues, with (over)expression being detected by quantitative PCR. Analysis of prostatic cancer samples showed that CD24 was upregulated in many of them. Sections of tissue, isolated from prostatic cancer patients, or subjects at risk, were incubated sequentially with anti-human CD4 murine monoclonal antibodies; biotinylated second antibody; streptavidin-conjugated horseradish peroxidase and then diaminobenzidine as colour former (brown). The samples were counterstained with hemalum (blue). Malignant cells stained strongly but non-malignant cells only weakly. In 15 of 63 samples of adenocarcinoma, membrane and cytoplasmic staining was very strong, and lymph node metastases were also stained. ADR65805-ADR66954 represent the polynucleotide and polypeptide sequences used in the method of the invention.

XX SQ Sequence 371 AA;

Query Match 54.1%; Score 59; DB 8; Length 371;

Best Local Similarity 50.0%; Pred. No. 1.1;

Matches 10; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 EGYIQNYRCRGDDSKVQEAR 20

DB 232 DGYIEYICRGNAERVKRGR 251

RESULT 25

ADS24124

ID ADS24124 standard; protein; 847 AA.

XX AC ADS24124;

XX DT 02-DEC-2004 (first entry)

XX DE Bacterial polypeptide #13157.

XX KW Recombinant DNA construct; transformed plant; improved plant property; cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis; pathogen tolerance; pest tolerance; plant disease resistance; cell cycle pathway modification; plant growth regulator; homologous recombination; seed oil yield; protein yield; carbohydrate; nitrogen; phosphorus; photosynthesis; lignin; galactomannan; bacterial polypeptide.

XX OS Bacteria.

XX XN US2003233675-A1.

XX PD 18-DEC-2003.

XX PF 20-FEB-2003; 2003US-00369493.

XX PR 21-FEB-2002; 2002US-0360039P.

XX PA (CAOY/) CAO Y.

XX PA (HINK/) HINKLE G J.

XX PA (SLAT/) SLATER S C.

XX PA (CHEN/) CHEN X.

XX PA (GOLD/) GOLDMAN B S.

XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX DR WPI; 2004-061375/06.

XX PT New recombinant DNA construct comprising a promoter positioned to provide for expression of a polynucleotide encoding a polypeptide from a microbial source, useful for producing plants with improved properties.

XX PS Claim 1; SEQ ID NO 13157; 122pp; English.

XX CC The invention relates to a recombinant DNA construct comprising a promoter functional in a plant cell, where the promoter is positioned to provide for expression of a polynucleotide encoding a polypeptide from a microbial source. The invention also relates to a transformed plant comprising the recombinant DNA construct and a method of producing a transformed plant having an improved property. The plant is a crop plant such as maize or soybean. The method of producing a transformed plant having an improved property comprises transforming a plant with the recombinant DNA construct and growing the transformed plant, where the polynucleotide or polypeptide is useful for improving plant properties. The recombinant DNA construct is useful for producing plants with improved plant properties, e.g. improved cold, heat or drought tolerance, increased resistance to plant disease, better growth rate by modification of the cell cycle pathway with plant growth regulators, increased rate of homologous recombination, modified seed oil or protein yield and/or content, improved yield by modification of carbohydrate, nitrogen or phosphorus use and/or uptake, by modification of photosynthesis or by providing improved plant growth and development under at least one stress condition, improved lignin production or improved galactomannan production. This sequence represents a bacterial polypeptide used in the scope of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 847 AA;

Query Match 45.0%; Score 49; DB 8; Length 847;

Best Local Similarity 53.3%; Pred. No. 1e+02;

Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 GYIQNYRCRGDDSKV 16

DB 181 GFIOKRCRGQDIEI 195

RESULT 26

ADS24139

ID ADS24139 standard; protein; 1327 AA.

XX AC ADS24139;

XX DT 02-DEC-2004 (first entry)

XX DE Bacterial polypeptide #13172.

XX KW Recombinant DNA construct; transformed plant; improved plant property; cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis; pathogen tolerance; pest tolerance; plant disease resistance; cell cycle pathway modification; plant growth regulator; homologous recombination; seed oil yield; protein yield; carbohydrate; nitrogen; phosphorus; photosynthesis; lignin; galactomannan; bacterial polypeptide.

XX OS Bacteria.

XX XN US2003233675-A1.

XX PD 18-DEC-2003.

XX PF 20-FEB-2003; 2003US-00369493.

PR 21-FEB-2002; 2002US-0360039P.
XX (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
DR
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 13172; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 1327 AA;
Query Match 45.0%; Score 49; DB 8; Length 1327;
Best Local Similarity 53.3%; Pred. No. 1.6e+02;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 2 GYIQNYRCRGDDSKV 16
DB 255 GFQKRCRGQDIEI 269
RESULT 27
ADB74255
ID ADB74255 standard; protein; 705 AA.
XX
AC ADB74255;
XX
DT 04-DEC-2003 (first entry)
XX
DE Mycobacterium tuberculosis non-naturally occurring peptide #3.
XX
KW Non-naturally occurring peptide; anion pump protein; tuberculosis;
KW hypersensitivity reaction; tuberculostatic.
XX
OS Mycobacterium tuberculosis.
XX
PN US6583266-B1.
XX
PD 24-JUN-2003.
XX

PF 16-SEP-1994; 94US-00311731.
XX
PR 19-AUG-1993; 93US-00109181.
PR 22-OCT-1993; 93US-00142558.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Smith DR, Mao J;
XX WPI; 2003-656441/62.
DR N-PSDB; ADB74252.
XX
PT New Mycobacterium tuberculosis anion pump peptide useful for as
PT tuberculosis vaccine and diagnosis of tuberculosis infection.
XX
XX Disclosure; SEQ ID NO 4; 26pp; English.
XX
CC The invention relates to a non-naturally occurring peptide of
CC Mycobacterium tuberculosis comprising an amino acid sequence
CC corresponding to an anion pump protein. The invention also relates to a
CC non-naturally occurring nucleic acid corresponding to a DNA sequence of
CC Mycobacterium tuberculosis or Mycobacterium leprae. The new peptide is
CC useful as a vaccine against Mycobacterium tuberculosis or Mycobacterium
CC leprae or for screening for new tuberculosis drugs. Purified proteins
CC derived from the sequences of the invention may elicit a specific immune
CC response. The peptide may also be used to detect hypersensitivity
CC reactions of individuals exposed to Mycobacterium tuberculosis or
CC Mycobacterium leprae. The proteins and peptides may be affixed to solid
CC supports to detect antibodies typical of hypersensitivity reactions, from
CC a patient's sera. This sequence represents Mycobacterium tuberculosis non
CC -naturally occurring peptide of the invention. Note: The sequence data
CC for this patent did not form part of the printed specification but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 705 AA;
Query Match 42.7%; Score 46.5; DB 7; Length 705;
Best Local Similarity 50.0%; Pred. No. 2.1e+02;
Matches 8; Conservative 5; Mismatches 2; Indels 1; Gaps 1;
QY 1 EGYIQNYRCRGDDSKV 16
DB 378 DGWV-TYRCRADTTEV 392
RESULT 28
ABB62735
ID ABB62735 standard; protein; 340 AA.
XX
AC ABB62735;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 14997.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE) PE CORP NY.
PA
XX Venter JC, Adams M, Li PWD, Myers EW;
PI

Sequences AAU39105-AAU68017 represent *Propionibacterium acnes* immunogenic polypeptides. The proteins and their associated DNA sequences are used in the treatment, prevention and diagnosis of medical conditions caused by *P. acnes*. The disorders include SAPHO syndrome (synovitis, acne, pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis. *P. acnes* is also involved in infections of bone, joints and the central nervous system, however it is particularly involved in the inflammatory lesions associated with acne vulgaris. A method for detecting the presence or absence of *P. acnes* in a patient comprises contacting a sample with a binding agent that binds to the proteins of the invention and determining the amount of bound protein in the sample. The polypeptides may be used as antigens in the production of antibodies specific for *P. acnes* proteins. These antibodies can be used to downregulate expression and activity of *P. acnes* polypeptides and therefore treat *P. acnes* infections. The antibodies may also be used as diagnostic agents for determining *P. acnes* presence, for example, by enzyme linked immunosorbent assay (ELISA). Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences)

xx SQ Sequence 84 AA;

Query Match 41.3%; Score 45; DB 4; Length 84;
Best Local Similarity 37.5%; Pred. No. 37;
Matches 6; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 5 QNYRCRGDDSKVOEAR 20
 : : ||| : : : : : : :
Db 35 RHVACRGEGVRIQDSR 50

RESULT 30
ABM56755
ID ABM56755 standard; protein; 84 AA.
XX AC ABM56755;
XX DT 20-OCT-2003 (first entry)
XX DE *Propionibacterium acnes* predicted ORF-encoded polypeptide #21431.
XX ACne vulgaris; antiseborrheic; dermatological; antibacterial;
XX KW immunostimulant; immune response; vaccine.
XX OS *Propionibacterium acnes*.
XX PN WO2003033515-A1.
XX PD 24-APR-2003.
XX PF 11-OCT-2002; 2002WO-US032727.
XX PR 15-OCT-2001; 2001US-00978825.
XX (CORI-) CORIXA CORP.
XX PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MU, Benson DR, Jones R, Carter D;
PI Barth B, Valliave-Douglas J;
XX WPI; 2003-381789/36.
DR N-PSDB; ACF64537.
XX New *Propionibacterium acnes* polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a *P. acnes* protein.
XX
PS Example 1; SEQ ID NO 21431; 1481pp; English.
XX
XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733) encoding a *Propionibacterium acnes* protein. The invention also relates to polynucleotides encoded by the polynucleotides (ABM35624-ABM64536) and to

CC immunogenic fragments of *P. acnes* polypeptides. The invention
 CC additionally encompasses expression vectors and host cells comprising a
 CC polynucleotide of the invention; antibodies against polypeptides of the
 CC invention; fusion proteins comprising a polypeptide of the invention; a
 CC method for stimulating an immune response specific for a *P. acnes*
 CC polypeptide and an isolated T cell population comprising T cells prepared
 CC via this method; a vaccine composition (comprising *P. acnes* polypeptides,
 CC polynucleotides, antibodies, fusion proteins, T cell populations, or
 CC antigen-presenting cells that express the polypeptide); a method and kit
 CC for detecting or determining the presence or absence of *P. acnes* in a
 CC patient; and a method for inhibiting the development of *P. acnes* in a
 CC patient. The *P. acnes* polypeptides, polynucleotides, antibodies, fusion
 CC proteins, T cell populations or antigen-presenting cells that express the
 CC polypeptides are useful for diagnosing, preventing or treating acne
 CC vulgaris, or for stimulating an immune response specific for a *P. acnes*
 CC protein. The polynucleotides can also be used as probes or primers for
 CC nucleic acid hybridisation. The vaccine composition is useful for the
 CC stimulation of an immune response against *P. acnes*, or for treating acne,
 CC and the kit is useful for performing a diagnostic assay. The present
 CC sequence represents a polypeptide predicted to be encoded by an ORF (open
 CC reading frame) contained within the *P. acnes* polynucleotides of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 84 AA;

Query Match 41.3%; Score 45; DB 6; Length 84;
 Best Local Similarity 37.5%; Pred. No. 37;
 Matches 6; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 5 QNYRCRGDDSKVQEAR 20
 :|||::: :|||:::
 Db 35 RHYACGEGVRQDSR 50

RESULT 31

ADL92508
 ID ADL92508 standard; protein; 135 AA.

XX AC ADL92508;

XX DT 20-MAY-2004 (first entry)

XX DE iMab111 VAP protein fragment.

XX KW immunoglobulin superfamily; beta-barrel; cytostatic; antimicrobial;
 cancer; microbial infection; protein-coordinate data.

XX OS Unidentified.

XX PN US2003215914-A1.

XX PD 20-NOV-2003.

XX PF 10-DEC-2002; 2002US-00316194.

XX PR 10-DEC-2001; 2001US-00016516.

XX PA (HOUT/) HOUTZAGER E.

XX PA (VIJN/) VIJN I M C.

XX PA (SIJMN/) SIJMONS P C.

XX PI Houtzager E, Vijjn IMC, Sijmons PC;

XX DR WPI; 2004-021843/02.

XX PT New proteinaceous molecule comprising a binding peptide and a core
 PT comprising a beta-barrel, useful for preparing a composition for
 PT diagnosing or treating a condition involving unwanted proteins, cells or
 PT microorganisms e.g., cancer.

XX PS Example 3; Page 30; 84pp; English.

XX This invention describes a novel proteinaceous molecule (e.g. from an
 CC immunoglobulin superfamily origin) which comprises a binding peptide and
 CC a core comprising a beta-barrel of at least 4 strands. The invention also
 CC describes a method for identifying the proteinaceous molecule, cells
 CC comprising the isolated proteinaceous molecule, a method for producing a
 CC nucleic acid encoding the proteinaceous molecule capable of displaying at
 CC least one desired peptide sequence, displaying a desired peptide
 CC sequence, producing a library including artificial binding peptides,
 CC separating a substance from a mixture, a pharmaceutical composition
 CC comprising the proteinaceous molecule for treating a pathological
 CC condition involving unwanted proteins, cells or microorganisms, a
 CC diagnostic assay comprising the isolated proteinaceous molecule, a gene
 CC delivery vehicle comprising a nucleic acid encoding the isolated
 CC proteinaceous molecule and a nucleic acid encoding a gene of interest and
 CC a chromatography column comprising the proteinaceous molecule and a
 CC packaging material. The alteration of the molecule comprises a post-
 CC translational modification. The products of the invention have cytostatic
 CC and antimicrobial activity. The proteinaceous molecule is useful for
 CC preparing a composition for diagnosing or treating a pathological
 CC condition involving unwanted proteins, cells or microorganisms, e.g.,
 CC cancer or microbial infections.

XX SQ Sequence 135 AA;

Query Match 41.3%; Score 45; DB 8; Length 135;
 Best Local Similarity 53.3%; Pred. No. 61;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 GYIQNYRCRGDDSKV 16
 |||:::|:::
 Db 119 GYDSHYRCQGTDTV 133

RESULT 32

ADR41823
 ID ADR41823 standard; protein; 135 AA.

XX AC ADR41823;

XX DT 04-NOV-2004 (first entry)

XX DE Versatile affinity protein iMab111.

XX KW cosmetic; target; softener; hair; perfume; deodorant; mouth wash;
 KW cleaning; hair dye; lipstick; rouge; skin colouring; affinity region;
 KW versatile affinity protein; VAP; protein co-ordinate data.

XX OS Lama glama.

XX PN WO2004069211-A2.

XX PD 19-AUG-2004.

XX PF 10-DEC-2003; 2003WO-NL000876.

XX PR 10-DEC-2002; 2002EP-00080206.

XX PR 10-DEC-2002; 2002US-0432906P.

XX PA (LMAB-) L-MABS BV.

XX PI Houtzager E, Vijjn IMC, Sijmons PC, Valinotti T, Mudge G, Padel A;

XX DR WPI; 2004-604350/58.

XX PT Applying cosmetic substance to desired target molecule, involves
 PT providing conjugate of proteinaceous substance having specific affinity
 PT for target molecule linked to cosmetic substance.

XX PS Claim 18; Fig 22; 213pp; English.

XX The invention relates to a novel method for applying a cosmetic substance
 CC to a desired target molecule. The method comprises providing a conjugate

CC of a proteinaceous substance having specific affinity for the target
 CC molecule linked to a cosmetic substance, where the resulting connection
 CC between cosmetic substance and target molecule can be disrupted upon the
 CC presence of a chemical and/or physical signal. The invention further
 CC comprises: a conjugate obtained by the method; a cosmetic composition
 CC comprising the conjugate; a detergent and/or softener composition
 CC comprising the conjugate; and a di- or multivalent proteinaceous
 CC substance having specific affinity for two or more target molecules
 CC present in hair, where epitopes are recognised on two or more target
 CC molecules may be the same or different. The method is useful for applying
 CC a cosmetic substance to a desired target molecule. The di- or multivalent
 CC proteinaceous substance is useful for a perfume, a deodorant, a mouth
 CC wash, cleaning, hair dye, a lipstick, rouge or another skin colouring
 CC composition. This sequence represents a versatile affinity protein (VAP)
 CC of the invention.

XX

SQ Sequence 135 AA;

Query Match 41.3%; Score 45; DB 8; Length 135;
 Best Local Similarity 53.3%; Pred. No. 61;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 GYIQNYRCRGDDSKV 16
 || :|||:|
 Db 119 GYDSHYRCQGTDTV 133

RESULT 33

AAW99087

ID AAW99087 standard; protein; 822 AA.

XX

AC AAW99087;

XX

DT 13-MAY-1999 (first entry)

XX

DE Human serine protease BSSP-3.

XX

KW Serine protease; BSSP-3; brain tissue.

XX

OS Homo sapiens.

XX

PN WO9905290-A1.

XX

PD 04-FEB-1999.

XX

PF 24-JUL-1998; 98WO-JP003324.

XX

PR 24-JUL-1997; 97JP-00213969.

XX

PA (SUNR) SUNTORY LTD.

XX

PI Tsuruoka N, Yamashiro K, Yamaguchi N;

XX

DR WPI: 1999-142942/12.

XX

DR N-PSDB; AAX19024.

XX

PT New serine protease expressed in brain tissue - used in screening for

XX

PS potential serine protease inhibitors for drug use.

XX

PS Claim 1; Page 61-65; 69pp; Japanese.

XX

CC The present sequence is a serine protease designated BSSP-3, which is
 CC isolated from human brain tissue. Transformants may be used to produce
 CC the enzyme or its partial sequences. Products from the present invention
 CC are used for screening for potential peptide or non-peptide serine
 CC protease inhibitors or expression regulators for use as drugs

XX

SQ Sequence 822 AA;

Query Match

Best Local Similarity 41.3%; Score 45; DB 2; Length 822;

Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 YIQNYRCRGDDSKV 16
 || :|||:|
 Db 233 YMSNVRCRGDEENI 246

RESULT 35

QY 3 YIQNYRCRGDDSKV 16
 || :|||:|
 Db 180 YMSNVRCRGDEENI 193

RESULT 34

AAW83361

ID AAW83361 standard; protein; 875 AA.

XX

AC AAW83361;

XX

DT 17-FEB-1999 (first entry)

XX

DE Human neurotropsin.

XX

KW Human; neurotropsin; tumour inhibition; neurological disease;
 KW lung disease; gene therapy; drug development; stroke; brain injury;
 KW neurodegeneration; neuroinflammatory disease; multiple sclerosis;
 KW epilepsy; hypoxia; ischaemia; nerve transection; neovascularisation;
 KW emphysema; bronchitis.

XX

OS Homo sapiens.

XX

PN WO9849322-A1.

XX

PD 05-NOV-1998.

XX

PF 24-APR-1998; 98WO-IB000625.

XX

PR 26-APR-1997; 97CH-00000966.

XX

PA (SOND/) SONDEREGGER P.

XX

PI Sonderegger P;

XX

DR WPI: 1999-009438/01.

XX

DR N-PSDB; AAV72589.

XX

PT New human and murine neurotropsin - used, e.g. for inhibiting tumours,
 PT treatment of neurological or lung disease, including by gene therapy and
 PT in drug development.

XX

PS Claim 1; Page 20-24; 50pp; English.

XX

CC The present sequence represents human neurotropsin. Neurotropsin proteins
 CC and polynucleotides can be used: (i) to inhibit tumours, including
 CC metastases, e.g. of brain or retina; (ii) to minimise tissue damage
 CC caused by stroke or brain injury (having a protective effect on the
 CC penumbra zone); (iii) to treat or prevent neurodegeneration,
 CC neuroinflammatory disease (e.g. multiple sclerosis) or epilepsy; (iv) to
 CC increase survival of damaged neurons (e.g. in cases of hypoxia,
 CC ischaemia, nerve transection) and to stimulate regeneration and/or
 CC restoration of synapses; (v) to treat or prevent retinal disorders (e.g.
 CC degeneration or neovascularisation); (vi) to prevent apoptosis (or other
 CC causes of cell death) in the nervous system; (vii) to regenerate brain
 CC and/or nervous tissue; (viii) to treat pain; (ix) to improve brain
 CC performance, including learning and memory; (x) to treat or prevent a
 CC wide range of psychiatric disorders; and (xi) to treat brain or lung
 CC injury associated with protease expression (specifically emphysema or
 CC bronchitis)

XX

SQ Sequence 875 AA;

Query Match 41.3%; Score 45; DB 2; Length 875;

Best Local Similarity 50.0%; Pred. No. 4.5e+02;

Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 YIQNYRCRGDDSKV 16

|| :|||:|

Db 233 YMSNVRCRGDEENI 246

ADQ21226
 ID ADQ21226 standard; protein; 875 AA.
 AC ADQ21226;
 XX
 DT 26-AUG-2004 (first entry)
 XX
 DE Human soft tissue sarcoma-upregulated protein - SEQ ID 4046.
 XX
 KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.
 XX
 OS Homo sapiens.
 XX
 PN WO2004048938-A2.
 XX
 PD 10-JUN-2004.
 XX
 PF 26-NOV-2003; 2003WO-US038193.
 XX
 PR 26-NOV-2002; 2002US-0429739P.
 XX
 PA (PROT-) PROTEIN DESIGN LABS INC.
 XX
 PI Aziz N, Ginsburg WM, Zlotnik A;
 XX
 DR WPI; 2004-441208/41.
 XX
 PT Early detection of soft tissue sarcoma comprises determining expression
 of a gene in a first soft tissue sample and a normal soft tissue sample
 PT and comparing the gene expression, also useful in treating soft tissue
 PT sarcoma.
 XX
 PS Example 2; SEQ ID NO 4046; 210pp; English.
 XX
 CC The invention relates to a novel method for detecting soft tissue sarcoma
 CC which comprises obtaining a first soft tissue sample from an individual
 CC and a normal soft tissue sample from the same or different individual,
 CC determining the expression of a gene in both samples and comparing the
 CC expression of the gene in both soft tissue samples, where a higher level
 CC of protein expression in the first soft tissue sample indicates the
 CC presence of soft tissue sarcoma. The method of the invention has
 CC cytostatic applications and may be useful for detecting soft tissue
 CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
 CC acid sequences may be useful in diagnostic and screening applications.
 CC The current sequence is that of a human soft tissue sarcoma-upregulated
 CC protein of the invention. The current sequence is not shown within the
 CC specification per se but was submitted in CD format by the inventor.
 XX
 SQ Sequence 875 AA;
 Query Match 41.3%; Score 45; DB 8; Length 875;
 Best Local Similarity 50.0%; Pred. No. 4.5e+02; Indels 0; Gaps 0;
 Matches 7; Conservative 2; Mismatches 5;
 QY 3 YIQNYRCRGDDSKV 16
 Db 233 YWSNVRCRGDEENI 246
 DE
 RESULT 36
 ADQ89082
 ID ADQ89082 standard; protein; 875 AA.
 XX
 AC ADQ89082;
 XX
 DT 21-OCT-2004 (first entry)
 XX
 DE Human urological disorder related protein 14063 SEQ:34.
 XX
 KW urological disorder; uropathic; cytostatic; urinary incontinence;
 KW benign prostatic hyperplasia; human.
 XX
 OS Homo sapiens.

XX
 PN WO2004065576-A2.
 XX
 PD 05-AUG-2004.
 XX
 PF 14-JAN-2004; 2004WO-US000750.
 XX
 PR 15-JAN-2003; 2003US-0440318P.
 PR 04-FEB-2003; 2003US-0444783P.
 PR 27-MAR-2003; 2003US-0457901P.
 PR 08-MAY-2003; 2003US-0468775P.
 PR 19-MAY-2003; 2003US-0471614P.
 PR 16-JUN-2003; 2003US-0478742P.
 PR 18-JUL-2003; 2003US-0488529P.
 PR 30-JUL-2003; 2003US-0491156P.
 PR 02-SEP-2003; 2003US-0499594P.
 PR 26-SEP-2003; 2003US-0506332P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Karicheti V, Silos-Santiago I, Eliasof SD;
 XX
 DR WPI; 2004-562167/54.
 DR N-PSDB; ADQ89081.
 XX
 PT Use of polypeptides related to urological disorders, e.g. 44390, 54181,
 211 or for identifying a compound capable of treating a urological
 PT disorder or identifying and treating a subject having a urological
 PT disorder.
 XX
 PS Claim 1; SEQ ID NO 34; 542pp; English.
 XX
 CC The present invention describes the use of polypeptides related to
 CC urological disorders for identifying a compound capable of treating a
 CC urological disorder, identifying a subject having a urological disorder,
 CC or treating a subject having a urological disorder. Also described: (1) a
 CC method for identifying a compound capable of treating a urological
 CC disorder; (2) a method for identifying a subject having a urological
 CC disorder; and (3) a method for treating a subject having a urological
 CC disorder. The compound has uropathic and cytostatic activities. The
 CC polypeptides related to urological disorders are useful for identifying a
 CC compound capable of treating a urological disorder, identifying a subject
 CC having a urological disorder, or treating a subject having a urological
 CC disorder. Disorders include urinary incontinence and benign prostatic
 CC hyperplasia. The present sequence represents a human urological disorder
 CC related protein, which is used in the exemplification of the present
 CC invention.
 XX
 SQ Sequence 875 AA;
 Query Match 41.3%; Score 45; DB 8; Length 875;
 Best Local Similarity 50.0%; Pred. No. 4.5e+02; Indels 0; Gaps 0;
 Matches 7; Conservative 2; Mismatches 5;
 QY 3 YIQNYRCRGDDSKV 16
 Db 233 YWSNVRCRGDEENI 246
 DE
 RESULT 37
 ABUS4593
 ID ABUS4593 standard; protein; 191 AA.
 XX
 AC ABUS4593;
 XX
 DT 03-JUN-2003 (first entry)
 XX
 DE Human NOVX polypeptide #52.
 XX
 KW Human; NOVX; metabolic disorder; cardiomyopathy; diabetes; ASD;
 KW hypertension; congenital heart defect; aortic stenosis; valve disease;
 KW atrial septal defect; atrioventricular canal defect; ductus arteriosus;
 KW pulmonary stenosis; subaortic stenosis; ventricular septal defect; VSD;

KW tuberosus sclerosis; scleroderma; atherosclerosis; infectious disease;
KW obesity; anorexia; neurodegenerative disorder; Alzheimer's disease;
KW Parkinson's disease; immune disorder; haematopoietic disorder;
KW haemophilia; hypercoagulation; Crohn's disease; cancer.
XX Homo sapiens.
OS
XX
PN W0200281498-A2.
XX
PD 17-OCT-2002.
XX
PF 03-APR-2002; 2002WO-US010780.
XX
PR 03-APR-2001; 2001US-0281086P.
PR 03-APR-2001; 2001US-0281136P.
PR 05-APR-2001; 2001US-0281863P.
PR 05-APR-2001; 2001US-0281906P.
PR 06-APR-2001; 2001US-0282020P.
PR 10-APR-2001; 2001US-0282930P.
PR 10-APR-2001; 2001US-0282934P.
PR 12-APR-2001; 2001US-0283512P.
PR 13-APR-2001; 2001US-0283710P.
PR 17-APR-2001; 2001US-0284234P.
PR 19-APR-2001; 2001US-0285325P.
PR 20-APR-2001; 2001US-0285381P.
PR 20-APR-2001; 2001US-0285609P.
PR 23-APR-2001; 2001US-0285748P.
PR 23-APR-2001; 2001US-0285890P.
PR 24-APR-2001; 2001US-0286068P.
PR 25-APR-2001; 2001US-0286292P.
PR 27-APR-2001; 2001US-0287213P.
PR 02-MAY-2001; 2001US-0288257P.
PR 29-MAY-2001; 2001US-0294164P.
PR 30-MAY-2001; 2001US-0294484P.
PR 18-JUN-2001; 2001US-0298952P.
PR 19-JUN-2001; 2001US-0299237P.
PR 19-JUN-2001; 2001US-0299276P.
PR 12-SEP-2001; 2001US-0318750P.
PR 25-SEP-2001; 2001US-0324800P.
PR 25-SEP-2001; 2001US-0324802P.
PR 27-SEP-2001; 2001US-0325684P.
PR 17-OCT-2001; 2001US-0330143P.
PR 14-NOV-2001; 2001US-0332131P.
PR 14-NOV-2001; 2001US-0332240P.
PR 14-NOV-2001; 2001US-0332779P.
PR 21-NOV-2001; 2001US-0332115P.
PR 04-DEC-2001; 2001US-0337621P.
PR 03-JAN-2002; 2002US-0345783P.
PR 16-JAN-2002; 2002US-0350251P.
PR 02-APR-2002; 2002US-00114270.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Guo X, Kekuda R, Miller CE, Malyankar UM, Spytek KA;
PI Pattarajan M, Liu X, Gusev VY, Li L, Vernet CAM, Zerhusen BP;
PI Gorman L, Shency SG, Pena CEA, Smithson G, Burgess CE, Gerlach V;
PI Padigar M, Shinkets RA, Gangilli EA, Taupier RJ, Casman SJ, Ji W;
PI Anderson DW, Leite MW, Rastelli L, Edinger SR, Stone DJ;
PI Macdougall JR, Rothenberg MB, Mazur A, Millet I, Peyman JA;
PI Ellerman K;
XX
DR WPI; 2003-046858/04.
DR N-PSDB; ABX72221.
XX
XX New isolated NOVX polypeptide useful for treating atherosclerosis,
PT metabolic disorders, diabetes, obesity, infectious disease, anorexia,
PT neurodegenerative disorders, Alzheimer's disease and cancer.
XX
PS Claim 1; Page 205; 666pp; English.
XX
XX The invention relates to human polypeptides, termed NOVX, and the
CC polynucleotides encoding them. The polypeptides and polynucleotides are
CC useful for diagnosing disease, and screening for potential therapeutic

CC agents. The sequences are useful for treating metabolic disorders,
CC cardiomyopathy, diabetes, hypertension, congenital heart defects, aortic
CC stenosis, atrial septal defect (ASD), atrioventricular canal defect,
CC ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular
CC septal defect (VSD), valve diseases, tuborous sclerosis, scleroderma,
CC atherosclerosis, obesity, infectious disease, anorexia, neurodegenerative
CC disorders, Alzheimer's disease, Parkinson's disease, immune disorders,
CC haematopoietic disorders, haemophilia, hypercoagulation, Crohn's disease
CC and cancer. Sequences ABU54542-ABU54647 represent human NOVX polypeptides
XX of the invention
XX
SQ Sequence 191 AA;
Query Match 40.4%; Score 44; DB 6; Length 191;
Best Local Similarity 43.8%; Pred. No. 1.3e+02;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QY 4 IONYRCRGDDSKVQEA 19
Db 137 LQLYRCREDAAIYQA 152
RESULT 38
ADE58277
ID ADE58277 standard; protein; 385 AA.
XX
AC ADE58277;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human Protein Q01459, SEQ ID NO 4148.
XX
KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
XX
PN W02003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PR (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
XX WPI; 2003-268312/26.
DR GENBANK; Q01459.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a

CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (SN1)), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a human protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 385 AA;
 Query Match 40.4%; Score 44; DB 7; Length 385;
 Best Local Similarity 88.9%; Pred. No. 2.7e+02;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 YIQNYRCRG 11
 |||||
 Db 342 YIQNYRLRG 350

RESULT 39

ADM04173
 ID ADM04173 standard; protein; 485 AA.

XX AC ADM04173;

XX DT 20-MAY-2004 (first entry)

XX DE Human protein of the invention SEQ ID NO:2859.

XX KW human; gene therapy; diagnostic marker; pharmaceutical.

XX OS Homo sapiens.

XX FN EP1347046-A1.

XX PD 24-SEP-2003.

XX PF 12-APR-2002; 2002EP-00008400.

XX PR 22-MAR-2002; 2002JP-00137785.

XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 PI Yamamoto J, Isono Y, Kio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;

XX DR WPI; 2003-723558/69.

XX DR N-PSDB; ADM01730.

XX PT New polynucleotides and polypeptides are useful in gene therapy, for
 PT developing a diagnostic marker or medicines for regulating their
 PT expression and activity, or as a target of gene therapy.

XX PS Claim 1; SEQ ID NO 2858; 305pp; English.

XX CC The invention relates to a novel human polynucleotide and the encoded
 CC polypeptide. A polynucleotide of the invention may have a use in gene
 CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
 CC as a primer for synthesizing the polynucleotide or as a probe for
 CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are
 CC useful in gene therapy, for developing a diagnostic marker or medicines
 CC for regulating their expression and activity, or as a target of gene
 CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
 CC are useful as pharmaceutical agents. The present sequence represents a

CC protein sequence of the invention.

XX SQ Sequence 485 AA;

Query Match 40.4%; Score 44; DB 7; Length 485;
 Best Local Similarity 53.8%; Pred. No. 3.4e+02;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 YIQNYRCRGDDSK 15
 |::|||
 Db 466 YLYNWRCPGTDS 478

RESULT 40

AAAY79209
 ID AAAY79209 standard; protein; 498 AA.

XX AC AAAY79209;

XX DT 19-JUN-2000 (first entry)

XX DE Human transferase TRNSFS-1.

XX KW Transferase; TRNSFS-1; human; antitumour; cancer;
 KW gastrointestinal disorder; developmental disorder; genetic disorder;
 KW neurological disorder; reproductive disorder; smooth muscle disorder;
 KW immunological disorder; inflammation; diagnosis; therapy;
 KW myristoyl CoA:protein N-myristoyltransferase.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Modified-site 64 /note= "potential O-phosphorylation"

FT Modified-site 68 /note= "potential O-phosphorylation"

FT Peptide 150..228 /note= "myristoyl-CoA protein N-myristoyltransferase signature"

FT Modified-site 166 /note= "potential O-phosphorylation"

FT Modified-site 220 /note= "potential O-phosphorylation"

FT Peptide 240..285 /note= "myristoyl-CoA protein N-myristoyltransferase signature"

FT Modified-site 240 /note= "potential O-phosphorylation"

FT Peptide 246..254 /note= "myristoyl-CoA protein N-myristoyltransferase signature"

FT Modified-site 258 /note= "potential O-phosphorylation"

FT Modified-site 270 /note= "potential O-phosphorylation"

FT Peptide 286..362 /note= "myristoyl-CoA protein N-myristoyltransferase signature"

FT Modified-site 295 /note= "potential O-phosphorylation"

FT Modified-site 319 /note= "potential N-glycosylation"

FT Modified-site 325 /note= "potential O-phosphorylation"

FT Modified-site 350 /note= "potential O-phosphorylation"

FT Modified-site 350 /note= "potential O-phosphorylation"

FT Modified-site 417 /note= "potential O-phosphorylation"

FT Peptide 436..488 /note= "myristoyl-CoA protein N-myristoyltransferase signature"

PT Modified-site 460 /note= "potential N-glycosylation"

PT Modified-site 462 /note= "potential O-phosphorylation"

PT Peptide 468...474 /note= "myristoyl-CoA protein N-myristoyltransferase signature"

PT Modified-site 488 /note= "potential O-phosphorylation"

PT Modified-site 490 /note= "potential O-phosphorylation"

XX WO200014251-A2.

XX 16-MAR-2000.

XX 09-SEP-1999; 99WO-US020989.

XX 10-SEP-1998; 98US-00150657.

PR 04-NOV-1998; 98US-00186779.

PR 11-MAY-1999; 99US-0133642P.

XX (INCY-) INCYTE PHARM INC.

XX Tang YT, Corley NC, Guegler KJ, Baughn MR, Lal P, Yue H; Hillman JL, Azimzai Y;

PI WPI; 2000-256996/22.

XX N-PSDB; AAZ94201.

DR Human transferase proteins useful for preventing, diagnosing and treating cancers and developmental, gastrointestinal, genetic, immunological, neurological, reproductive and smooth muscle disorders.

XX Claim 1; Page 78-79; 113pp; English.

XX The present sequence is that of human transferase TRNSFS-1, 1 of 15 claimed novel human transferase proteins of the invention (see AAY79209-23). The sequence was deduced from a cDNA clone (see AAZ94201) isolated from a caecal tissue library. It shows homology to human myristoyl-CoA protein N-myristoyltransferase. TRNSFS-1 is expressed in reproductive, nervous and smooth muscle tissues, especially those associated with cancer and inflammation. The new human transferase proteins and the polynucleotides encoding them can be used in the diagnosis, prevention and treatment of cancer, developmental disorders, gastrointestinal disorders, genetic disorders, immunological disorders, neurological disorders, reproductive disorders, and smooth muscle disorders. The polypeptides can also be used to raise antibodies, and to screen for agonists and antagonists of transferase activity

XX Sequence 498 AA;

Query Match 40.4%; Score 44; DB 3; Length 498;

Best Local Similarity 53.8%; Pred. No. 3.5e+02;

Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 3 YIQNYRCRGDDSK 15
|:|:|:|:|:

Db 479 YLYNWRCPGTDS 491

RESULT 41

ADN95772

ID ADN95772 standard; protein; 498 AA.

XX ADN95772;

XX 01-JUL-2004 (first entry)

XX Human BEC/LEC-related protein sequence SeqID696.

XX growth; differentiation; blood endothelial cell; BEC;

XX lymphatic endothelial cell; LEC; hereditary lymphoedema; VEGFR-3;

KW Lymphatic growth agent; VEGF-C; VEGF-D; antiangiogenic; cytostatic; vasotropic; antiinflammatory; gene therapy; endothelial cell disorder; inflammatory disease; cancer metastasis; lymphatic system; human.

XX Homo sapiens.

PN WO2003080640-A1.

XX 02-OCT-2003.

XX 07-MAR-2003; 2003WO-US006900.

PR 07-MAR-2002; 2002US-0363019P.

XX (LUDW-) LUDWIG INST CANCER RES.

PA (LICN) LICENTIA LTD.

XX Alitalo K, Makinen T, Petrova T, Saharinen P, Saharinen J;

XX WPI; 2003-876999/81.

DR N-PSDB; ADN95773.

XX Example 1; SEQ ID NO 696; 176pp; English.

XX This invention relates to a method of differentially modulating the growth or differentiation of blood endothelial cells (BEC) or lymphatic endothelial cells (LEC) comprises contacting endothelial cells with a composition comprising an agent that differentially modulates blood or lymphatic endothelial cells. Treating hereditary lymphoedema comprises identifying a human subject with lymphoedema and with a mutation in at least one allele of a gene encoding a LEC protein, where the mutation correlates with lymphoedema in human subjects, and with the proviso that the LEC protein is not VEGFR-3; and administering to the subject a composition comprising a lymphatic growth agent selected from VEGF-C or VEGF-D polypeptides and polynucleotides. The invention may be useful for the development of compounds with an antiangiogenic, cytostatic, vasotropic or antiinflammatory activity or for gene therapy. The method is useful in modulating the growth or differentiation of blood endothelial cells or lymphatic endothelial cells, in treating hereditary lymphoedema, in screening for an endothelial cell disorder or of a drug on endothelial cells. The agent is useful in manufacturing a medicament for the differential modulation of blood vessel endothelial cell or lymphatic vessel endothelial cell growth or differentiation. The lymphatic growth agent may also be used in manufacturing a medicament for the treatment of hereditary lymphoedema resulting from a mutation in a LEC gene or of other diseases involving the lymphatic vessels, such as various inflammatory diseases and cancer metastasis via the lymphatic system. The present sequence is that of a human LEC/BEC differentially expressed protein which is related to the method of the invention. Note: This sequence does not appear in the specification but was obtained by the indexer using the source data given in table 14 of the specification.

XX Sequence 498 AA;

Query Match 40.4%; Score 44; DB 7; Length 498;

Best Local Similarity 53.8%; Pred. No. 3.5e+02;

Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 3 YIQNYRCRGDDSK 15
|:|:|:|:|:

Db 479 YLYNWRCPGTDS 491

RESULT 42

ADP23807

ID ADP23807 standard; protein; 498 AA.

XX ADP23807;

XX 18-NOV-2004 (first entry)

XX PRO polypeptide SEQ ID NO:985.

XX PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;
KW osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;
KW antiasthmatic; hepatotropic; respiratory; gene therapy; immune system.
XX Unidentified.
OS
XX WO2004041170-A2.
XX
XX 21-MAY-2004.
XX
XX 30-OCT-2003; 2003WO-US034312..
XX
XX 01-NOV-2002; 2002US-0423394P.
XX
XX (GETH) GENENTECH INC.
XX
XX Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;
PI Wu TD;
XX
XX WPI: 2004-419628/39.
XX
XX N-PSDB; ADP23806.
XX
XX New PRO polypeptides and polynucleotides, useful for treating e.g.
PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated
PT renal disease, or demyelinating diseases of the central or peripheral
PT nervous system.
XX
XX Claim 7; SEQ ID NO 985; 2940pp; English.
XX
XX The invention relates to a novel isolated nucleic acid and the PRO
CC polypeptide encoded by it. A protein of the invention has
CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide
CC of the invention may have a use in gene therapy. The PRO polypeptide, its
CC agonist, antagonist, or antibody that specifically binds to the
CC polypeptide is useful for treating an immune related disorder such as
CC systemic lupus erythematous, rheumatoid arthritis, osteoarthritis,
CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
CC disease, a demyelinating disease of the central or peripheral nervous
CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
CC disease, infectious or autoimmune chronic active hepatitis, primary
CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
CC disease, an autoimmune or immune-mediated skin disease, a bullous skin
CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
CC disease, asthma, allergic rhinitis, atopic dermatitis, food
CC hypersensitivity, urticaria, an immunologic disease of the lung,
CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
CC pneumonitis, a transplantation associated disease, graft rejection or
CC graft-versus-host disease. The present sequence represents a PRO protein
CC of the invention.
XX
XX Sequence 498 AA;
SQ
Query Match 40.4%; Score 44; DB 8; Length 498;
Best Local Similarity 53.8%; Pred. No. 3.5e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 3 YLYNRCRGDSSK 15
|:|:|:|:|:
DB 479 YLYNRCRGDSE 491
RESULT 43
ADS24143
ID ADS24143 standard; protein; 1100 AA.
XX

AC ADS24143;
XX
XX 02-DEC-2004 (first entry)
XX
XX Bacterial polypeptide #13176.
DE
XX
XX Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
XX Bacteria.
XX
XX US2003233675-A1.
XX
XX 18-DEC-2003.
XX
XX 20-FEB-2003; 2003US-00369493.
XX
XX 21-FEB-2002; 2002US-0360039P.
XX
XX (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
PI WPI: 2004-061375/06.
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 13176; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
XX Sequence 1100 AA;
SQ
Query Match 40.4%; Score 44; DB 8; Length 1100;
Best Local Similarity 46.7%; Pred. No. 8.2e+02;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
QY 2 GYIQNYRCRGDSSK 16
|:|:|:|:|:
XX

```

Db          59  GFTQKRKCHQDIEI 73
RESULT 44
ABU62071
ID  ABU62071 standard; protein; 1907 AA.
XX
AC  ABU62071;
XX
DT  04-SEP-2003 (first entry)
XX
DE  Human skeletal muscle alpha-kinase (SK).
XX
KW  Human; mammalian kinase; melanoma alpha-kinase; MK; HK; KK; SK; IK;
KW  heart alpha-kinase; kidney alpha-kinase; skeletal muscle alpha-kinase;
KW  lymphocyte alpha-kinase; alpha-kinase catalytic domain;
XX  ion channel domain; enzyme.
XX
OS  Homo sapiens.
XX
PN  US2002177205-A1.
XX
PD  28-NOV-2002.
XX
PF  10-APR-2001; 2001US-00832292.
XX
PR  03-AUG-2000; 2000US-00632131.
XX
PA  (RYAZ/) RYAZANOV A.
XX
PI  Ryzanov A;
XX
DR  WPI; 2003-352609/33.
DR  N-PSDB; ACD26093.
XX
PT  Novel mammalian alpha-kinase proteins, including melanoma, kidney, heart,
PT  skeletal muscle or lymphocyte alpha-kinase protein useful for treating an
PT  animal in need of increased activity of mammalian alpha-kinase.
XX
PS  Claim 26; Fig 10B; 69pp; English.
XX
CC  The present invention relates to the isolation of novel mammalian kinases
CC  designated melanoma alpha-kinase (MK), heart alpha-kinase (HK), kidney
CC  alpha-kinase (KK), skeletal muscle alpha-kinase (SK), and lymphocyte
CC  alpha-kinase (LK), and the polynucleotide sequences encoding them. The
CC  alpha-kinase proteins are characterised by the presence of at least two
CC  domains, one being an alpha-kinase catalytic domain and other being an
CC  ion channel domain. The proteins are useful for treating an animal in
CC  need of increased activity of melanoma, kidney, heart, skeletal muscle or
CC  lymphocyte alpha-kinase. An antibody to the kinase is useful for
CC  detecting the presence of an melanoma, kidney, heart, skeletal muscle or
CC  lymphocyte alpha-kinase in a sample. ABU62065-ABU62072 represent the
CC  alpha-kinases of the invention
XX
SQ  Sequence 1907 AA;
Query Match          40.4%; Score 44; DB 6; Length 1907;
Best Local Similarity 43.8%; Pred. No. 1.5e+03;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY  4  IQNYRCRGDSKQVEA 19
:  |||||:|:|:|
Db  339  LQLYRCREDAAIYQA 354

RESULT 45
ABG23434
ID  ABG23434 standard; protein; 152 AA.
XX
AC  ABG23434;
XX
DT  18-FEB-2002 (first entry)
XX
DE  Novel human secretory protein, Seq ID No 610.
XX

```

```

DE  Novel human diagnostic protein #23425.
XX
KW  Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW  food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS  Homo sapiens.
XX
PN  WO200175067-A2.
XX
PD  11-OCT-2001.
XX
PF  30-MAR-2001; 2001WO-US008631.
XX
PR  31-MAR-2000; 2000US-00540217.
PR  23-AUG-2000; 2000US-00649167.
XX
PA  (HYSE-) HYSEQ INC.
XX
PI  Drmanac RT, Liu C, Tang YT;
XX
DR  WPI; 2001-639362/73.
DR  N-PSDB; AAS87621.
XX
PT  New isolated polynucleotide and encoded polypeptides, useful in
PT  diagnostics, forensics, gene mapping, identification of mutations
PT  responsible for genetic disorders or other traits and to assess
PT  biodiversity.
XX
PS  Claim 20; SEQ ID NO 53793; 103pp; English.
XX
CC  The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC  sequences. (I) is useful as hybridisation probes, polymerase chain
CC  reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC  and in recombinant production of (II). The polynucleotides are also used
CC  in diagnostics as expressed sequence tags for identifying expressed
CC  genes. (I) is useful in gene therapy techniques to restore normal
CC  activity of (II) or to treat disease states involving (II). (II) is
CC  useful for generating antibodies against it, detecting or quantitating a
CC  polypeptide in tissue, as molecular weight markers and as a food
CC  supplement. (II) and its binding partners are useful in medical imaging
CC  of sites expressing (II). (I) and (II) are useful for treating disorders
CC  involving aberrant protein expression or biological activity. The
CC  polypeptide and polynucleotide sequences have applications in
CC  diagnostics, forensics, gene mapping, identification of mutations
CC  responsible for genetic disorders or other traits to assess biodiversity
CC  and to produce other types of data and products dependent on DNA and
CC  amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC  amino acid sequences of the invention. Note: The sequence data for this
CC  patent did not appear in the printed specification, but was obtained in
CC  electronic format directly from WIPO at
CC  ftp.wipo.int/pub/published_pct_sequences
XX
SQ  Sequence 152 AA;
Query Match          39.4%; Score 43; DB 4; Length 152;
Best Local Similarity 58.3%; Pred. No. 1.4e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY  4  IQNYRCRGDSK 15
:  |||||:|
Db  66  LDNHRCRGDLTK 77

RESULT 46
AAU28253
ID  AAU28253 standard; protein; 237 AA.
XX
AC  AAU28253;
XX
DT  18-DEC-2001 (first entry)
XX
DE  Novel human secretory protein, Seq ID No 610.
XX

```

SQ	Sequence 237 AA;
Query Match	39.4%; Score 43; DB 4; Length 237;
Best Local Similarity	58.3%; Pred. No. 2.3e+02;
Matches	7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
Oy	4 IQNYRCRGDSDK 15 : : :
Dd	3 LDNHRCRGDLTK 14
RESULT 47	
ADS12199	
ID	ADS12199 standard; protein; 237 AA.
XX	ADS12199;
XX	AC AC
XX	16-DEC-2004 (first entry)
DT	XX
DE	Human therapeutic contig protein - SEQ ID 2436.
XX	
KW	antiinflammatory; neuroprotective; antianaemic; cytostatic; vulnerary; inflammatory; haematopoiesis; immunity; neurodegenerative; stem cell; aplastic anaemia; cancer; wound healing; gene therapy.
KW	Homo sapiens.
OS	
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 1..237 /label= Unknown, OTHER
FT	/note= "OTHER = In-frame STOP codon"
XX	
PX	WO2004080148-A2.
PN	
XX	
PD	23-SEP-2004.
XX	
PF	30-SEP-2003; 2003WO-US030720.
XX	
PR	02-OCT-2002; 2002US-0416186P.
XX	
PA	(NUVE-) NUVELO INC.
PI	Tang YT, Asundi V, Ren F, Zhang J, Wehrman T, Wang Z, Ma Y;
PI	Wang D, Chen R, Zhao QA, Wang J, Ghosh M, Xue AJ, Weng G, Zhou P;
XX	
DR	WPI: 2004-666857/65.
DR	N-PSDB; ADS11601.
XX	
PT	New polynucleotide, useful in preparing a composition for diagnosing or treating inflammatory, neurodegenerative or stem cell disorders, e.g., aplastic anemia or cancer for promoting wound healing.
XX	
PS	Example 2; SEQ ID NO 2436; 718pp; English.
CC	The invention relates to a novel isolated polynucleotide and the encoded polypeptide. The molecules of the invention demonstrate antiinflammatory, neuroprotective, antianaemic, cytostatic and vulnerary activities and may be useful in preparing a composition for diagnosing or treating inflammatory, haematopoietic, immune, neurodegenerative or stem cell disorders, such as aplastic anaemia or cancer, as well as for promoting wound healing. The molecules may also be utilised during gene therapy procedures. The current sequence is that of a human therapeutic contig protein of the invention.
XX	
SQ	Sequence 237 AA;
Query Match	39.4%; Score 43; DB 8; Length 237;
Best Local Similarity	58.3%; Pred. No. 2.3e+02;
Matches	7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
Oy	4 IQNYRCRGDSDK 15 : : :
Dd	3 LDNHRCRGDLTK 14

```

RESULT 48
ADN72921
ID   ADN72921 standard; protein; 371 AA.
AC   ADN72921;
XX
XX   15-JUL-2004 (first entry)
XX
XX   Thale cress protein upregulated in E2Fa/Dpa expressing plants SeqID 816.
XX
XX   plant; transgenic; E2Fa/Dpa transcription factor; growth regulator;
XX   animal feed product; thale cress; cell wall biosynthesis;
XX   nitrogen metabolism; carbon metabolism.
XX
XX   Arabidopsis thaliana.
XX
XX   WO2004035798-A2.
XX
XX   29-APR-2004.
XX
XX   20-OCT-2003; 2003WO-EP011658.
XX
XX   18-OCT-2002; 2002EP-00079408.
XX
XX   (CROP-) CROPDESIGN NV.
XX
XX   Inze D, De Veylder L, Vlieghe K;
XX
XX   WPI; 2004-348466/32.
XX
XX   N-PSDB; ADN72920.
XX
XX
XX   Altering plant characteristics, useful for producing plants for enzyme or
XX   pharmaceutical production comprises modifying in a plant, expression of
XX   one or more nucleic acids and/or modifying level or activity of one or
XX   more proteins.
XX
XX   Claim 1; SEQ ID NO 816; 134pp; English.
XX
XX   This invention relates to a novel method for altering one or more plant
XX   characteristics. Specifically, it refers to identifying genes that are up
XX   - or down-regulated in transgenic plants overexpressing the heterodimeric
XX   C2Fa/Dpa transcription factor of Arabidopsis and using these sequences to
XX   alter plant characteristics accordingly. The present invention describes
XX   generating transgenic plants for the production of growth regulators,
XX   enzymes, therapeutics, pharmaceuticals and animal feed products, where
XX   the altered plant characteristics are selected from increased yield or
XX   biomass, enhanced survival capacity, stress tolerance, plant architecture
XX   or physiology, altered endoreplication, biochemistry, signal
XX   transduction, storage lipid mobilisation and/or altered photosynthesis,
XX   each relative to the corresponding wild type plants. Accordingly, these
XX   sequences can also be useful as positive or negative selectable markers
XX   during transformation of cells or tissues. The identified genes play a
XX   role in a variety of biological processes such as DNA replication, cell
XX   wall biosynthesis, nitrogen and/or carbon metabolism or they function as
XX   transcription factors. This polypeptide sequence is thale cress protein
XX   expressed by a gene upregulated 1.3 fold or more in plants overexpressing
XX   the E2Fa/Dpa transcription factor, given in an exemplification of the
XX   invention.
XX
XX   Query Match          39.4%; Score 43; DB 8; Length 371;
XX   Best Local Similarity 47.1%; Pred. No. 3.7e+02;
XX   Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
XX
XX   Qy      4 IQNYRCRGDSDSKVQEAR 20
XX   :||| ||| ||| |||
XX   Db      301 VQNCRCDSDENVMWEFR 317
XX
XX
XX   RESULT 49

```

```

ADJ50947
ID   ADJ50947 standard; protein; 421 AA.
XX
XX   ADJ50947;
XX
XX   06-MAY-2004 (first entry)
XX
XX   Human novel protein NOV4b.
XX
XX   Human; NOVX; autoimmune disease; Alzheimer's disease; stroke; allergy;
XX   Parkinson's disease; Huntington's disease; multiple sclerosis; anxiety;
XX   pain; diabetes; graft versus host disease; pancreatitis; obesity; ulcer;
XX   anaemia; cancer; viral infection; bacterial infection;
XX   parasitic infection.
XX
XX   Homo sapiens.
XX
XX   US2004030096-A1.
XX
XX   12-FEB-2004.
XX
XX   01-AUG-2002; 2002US-00210281.
XX
XX   02-AUG-2001; 2001US-0309501P.
XX   03-AUG-2001; 2001US-0310291P.
XX   08-AUG-2001; 2001US-0310951P.
XX   09-AUG-2001; 2001US-0311292P.
XX   13-AUG-2001; 2001US-0311979P.
XX   14-AUG-2001; 2001US-0312203P.
XX   17-AUG-2001; 2001US-0313201P.
XX   20-AUG-2001; 2001US-0313643P.
XX   20-AUG-2001; 2001US-0313702P.
XX   21-AUG-2001; 2001US-0314031P.
XX   23-AUG-2001; 2001US-0314466P.
XX   28-AUG-2001; 2001US-0315403P.
XX   29-AUG-2001; 2001US-0315853P.
XX   05-MAR-2002; 2002US-0361775P.
XX   05-MAR-2002; 2002US-0361832P.
XX
XX   (GORM/) GORMAN L.
XX   (ZERR/) ZERHUSEN B D.
XX   (EDIN/) EDINGER S R.
XX   (PADI/) PADIGARU M.
XX   (GUOX/) GUO X.
XX   (KEKU/) KEKUDA R.
XX   (ZHON/) ZHONG M.
XX   (PATT/) PATTURAJAN M.
XX   (MILL/) MILLER C E.
XX   (JIWW/) JI W.
XX   (PENA/) PENA C E A.
XX   (BURG/) BURGESS C E.
XX   (SCIO/) SCIORE P.
XX   (STON/) STONE D J.
XX   (TAUP/) TAUPIER R J.
XX   (CASM/) CASMAN S J.
XX   (ROTH/) ROTHENBERG M E.
XX   (MALY/) MALYANKAR U M.
XX   (BOLD/) BOLDOG F L.
XX
XX   Gorman L, Zerhuseen BD, Edinger SR, Padigaru M, Guo X, Kekuda R;
XX   Zhong M, Patturajan M, Miller CE, Ji W, Pena CE, Burgess CE;
XX   Sciore P, Stone DJ, Taupier RJ, Casman SU, Rothenberg ME;
XX   Malyankar UM, Boldog FL;
XX
XX   WPI; 2004-168942/16.
XX   N-PSDB; ADJ50946.
XX
XX   New NOVX polypeptides and polynucleotides, useful in diagnosing, treating
XX   or preventing diseases or conditions, e.g. autoimmune disease,
XX   Alzheimer's disease, diabetes, graft versus host disease, cancer or viral
XX   or bacterial infections.
XX
XX   Claim 2; SEQ ID NO 12; 342pp; English.

```


CC parasitic infections (many more diseases and disorders are listed in the
 CC specification). The present sequence represents a NOVX protein.

XX
 SQ Sequence 424 AA;

Query Match 39.4%; Score 43; DB 8; Length 424;
 Best Local Similarity 58.3%; Pred. No. 4.3e+02;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 IQNYRCRGDDSK 15

Db 196 LDNHRCRGDLTK 207

Search completed: September 7, 2005, 19:54:42
 Job time : 167.186 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2005, 20:04:06 ; Search time 161 Seconds
(without alignments)
12.231 Million cell updates/sec

Title: US-10-812-238B-41

Perfect score: 32

Sequence: 1 CRGDD 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1774312 seqs, 393823214 residues

Total number of hits satisfying chosen parameters: 1774312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10F_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	100.0	6	17	US-10-812-238A-32
2	32	100.0	10	17	US-10-812-238A-20
3	32	100.0	13	17	US-10-812-238A-23
4	32	100.0	20	17	US-10-812-238A-2
5	32	100.0	43	16	US-10-437-963-157113
6	32	100.0	74	16	US-10-425-115-262985
7	32	100.0	77	16	US-10-437-963-170560
8	32	100.0	81	16	US-10-437-963-113416
9	32	100.0	152	15	US-10-389-566-938
10	32	100.0	155	16	US-10-425-115-261512
11	32	100.0	156	16	US-10-437-963-160624
12	32	100.0	156	16	US-10-425-115-306853
13	32	100.0	156	16	US-10-425-115-306853
14	32	100.0	156	16	US-10-425-115-306853
15	32	100.0	156	16	US-10-425-115-306853
16	32	100.0	156	16	US-10-425-115-306853
17	32	100.0	156	16	US-10-425-115-306853
18	32	100.0	156	16	US-10-425-115-306853
19	32	100.0	156	16	US-10-425-115-306853
20	32	100.0	156	16	US-10-425-115-306853
21	32	100.0	156	16	US-10-425-115-306853
22	32	100.0	156	16	US-10-425-115-306853
23	32	100.0	156	16	US-10-425-115-306853
24	32	100.0	156	16	US-10-425-115-306853
25	32	100.0	156	16	US-10-425-115-306853
26	32	100.0	156	16	US-10-425-115-306853
27	32	100.0	156	16	US-10-425-115-306853
28	32	100.0	156	16	US-10-425-115-306853
29	32	100.0	156	16	US-10-425-115-306853
30	32	100.0	156	16	US-10-425-115-306853
31	32	100.0	156	16	US-10-425-115-306853
32	32	100.0	156	16	US-10-425-115-306853
33	32	100.0	156	16	US-10-425-115-306853
34	32	100.0	156	16	US-10-425-115-306853
35	32	100.0	156	16	US-10-425-115-306853
36	32	100.0	156	16	US-10-425-115-306853
37	32	100.0	156	16	US-10-425-115-306853
38	32	100.0	156	16	US-10-425-115-306853
39	32	100.0	156	16	US-10-425-115-306853
40	32	100.0	156	16	US-10-425-115-306853
41	32	100.0	156	16	US-10-425-115-306853
42	32	100.0	156	16	US-10-425-115-306853
43	32	100.0	156	16	US-10-425-115-306853
44	32	100.0	156	16	US-10-425-115-306853
45	32	100.0	156	16	US-10-425-115-306853
46	32	100.0	156	16	US-10-425-115-306853
47	32	100.0	156	16	US-10-425-115-306853
48	32	100.0	156	16	US-10-425-115-306853
49	32	100.0	156	16	US-10-425-115-306853
50	32	100.0	156	16	US-10-425-115-306853
51	32	100.0	156	16	US-10-425-115-306853
52	32	100.0	156	16	US-10-425-115-306853
53	32	100.0	156	16	US-10-425-115-306853
54	32	100.0	156	16	US-10-425-115-306853
55	32	100.0	156	16	US-10-425-115-306853
56	32	100.0	156	16	US-10-425-115-306853
57	32	100.0	156	16	US-10-425-115-306853
58	32	100.0	156	16	US-10-425-115-306853
59	32	100.0	156	16	US-10-425-115-306853
60	32	100.0	156	16	US-10-425-115-306853
61	32	100.0	156	16	US-10-425-115-306853
62	32	100.0	156	16	US-10-425-115-306853
63	32	100.0	156	16	US-10-425-115-306853
64	32	100.0	156	16	US-10-425-115-306853
65	32	100.0	156	16	US-10-425-115-306853
66	32	100.0	156	16	US-10-425-115-306853
67	32	100.0	156	16	US-10-425-115-306853
68	32	100.0	156	16	US-10-425-115-306853
69	32	100.0	156	16	US-10-425-115-306853
70	32	100.0	156	16	US-10-425-115-306853
71	32	100.0	156	16	US-10-425-115-306853
72	32	100.0	156	16	US-10-425-115-306853
73	32	100.0	156	16	US-10-425-115-306853
74	32	100.0	156	16	US-10-425-115-306853
75	32	100.0	156	16	US-10-425-115-306853
76	32	100.0	156	16	US-10-425-115-306853
77	32	100.0	156	16	US-10-425-115-306853
78	32	100.0	156	16	US-10-425-115-306853
79	32	100.0	156	16	US-10-425-115-306853
80	32	100.0	156	16	US-10-425-115-306853
81	32	100.0	156	16	US-10-425-115-306853
82	32	100.0	156	16	US-10-425-115-306853
83	32	100.0	156	16	US-10-425-115-306853
84	32	100.0	156	16	US-10-425-115-306853

85	28	87.5	140	15	US-10-276-774-1857	Sequence 1857, Ap	158	27	84.4	57	16	US-10-425-115-190978	Sequence 190978,
86	28	87.5	147	15	US-10-424-599-180302	Sequence 180302,	159	27	84.4	72	16	US-10-425-115-315720	Sequence 315720,
87	28	87.5	158	15	US-10-424-599-162278	Sequence 162278,	160	27	84.4	72	15	US-10-424-599-223350	Sequence 223350,
88	28	87.5	160	15	US-10-085-198-96	Sequence 96, Appl	161	27	84.4	93	15	US-10-424-599-225782	Sequence 225782,
89	28	87.5	160	15	US-10-085-198-98	Sequence 98, Appl	162	27	84.4	105	16	US-10-425-115-251759	Sequence 251759,
90	28	87.5	160	16	US-10-729-473-7	Sequence 7, Appl	163	27	84.4	106	16	US-10-425-115-195840	Sequence 195840,
91	28	87.5	184	16	US-10-425-115-296818	Sequence 296818,	164	27	84.4	120	15	US-10-424-599-190435	Sequence 190435,
92	28	87.5	193	16	US-10-425-115-204748	Sequence 204748,	165	27	84.4	122	15	US-10-283-122A-62268	Sequence 62268, A
93	28	87.5	220	16	US-10-437-963-144344	Sequence 144344,	166	27	84.4	127	16	US-10-425-115-226461	Sequence 226461,
94	28	87.5	223	16	US-10-425-115-336315	Sequence 336315,	167	27	84.4	128	16	US-10-425-115-298115	Sequence 298115,
95	28	87.5	239	16	US-10-425-115-218457	Sequence 218457,	168	27	84.4	133	16	US-10-425-115-283941	Sequence 283941,
96	28	87.5	263	16	US-10-437-963-140042	Sequence 140042,	169	27	84.4	133	16	US-10-425-115-349300	Sequence 349300,
97	28	87.5	283	20	US-11-021-951-169	Sequence 169, App	170	27	84.4	148	15	US-10-424-599-161914	Sequence 161914,
98	28	87.5	294	16	US-10-425-115-331100	Sequence 331100,	171	27	84.4	158	16	US-10-437-963-109741	Sequence 109741,
99	28	87.5	323	9	US-09-764-898-181	Sequence 181, App	172	27	84.4	159	16	US-10-437-963-126010	Sequence 126010,
100	28	87.5	342	16	US-10-437-963-157655	Sequence 157655,	173	27	84.4	161	16	US-10-767-701-43187	Sequence 43187, A
101	28	87.5	348	16	US-10-739-930-8397	Sequence 8397, Ap	174	27	84.4	164	15	US-10-282-122A-43151	Sequence 43151, A
102	28	87.5	362	16	US-10-737-450-58	Sequence 58, Appl	175	27	84.4	186	16	US-10-767-701-32566	Sequence 32566, A
103	28	87.5	390	15	US-10-108-260A-4719	Sequence 4719, Ap	176	27	84.4	198	16	US-10-425-115-357665	Sequence 357665,
104	28	87.5	416	16	US-10-370-715B-150	Sequence 150, App	177	27	84.4	214	9	US-09-198-559-1	Sequence 1, Appl
105	28	87.5	417	15	US-10-453-478-18	Sequence 18, Appl	178	27	84.4	214	13	US-10-091-613-1	Sequence 1, Appl
106	28	87.5	421	15	US-10-368-438-10	Sequence 10, Appl	179	27	84.4	214	15	US-10-211-462-55	Sequence 55, Appl
107	28	87.5	435	8	US-08-459-455-51	Sequence 51, Appl	180	27	84.4	214	15	US-10-211-462-95	Sequence 95, Appl
108	28	87.5	435	9	US-09-354-697-9	Sequence 9, Appl	181	27	84.4	219	9	US-09-791-171-14	Sequence 14, Appl
109	28	87.5	435	14	US-09-851-873-97	Sequence 97, Appl	182	27	84.4	219	10	US-09-804-980-14	Sequence 14, Appl
110	28	87.5	435	14	US-10-280-670-9	Sequence 9, Appl	183	27	84.4	219	16	US-10-620-246-14	Sequence 14, Appl
111	28	87.5	435	18	US-10-973-858-16	Sequence 16, Appl	184	27	84.4	220	9	US-09-828-644-77	Sequence 77, Appl
112	28	87.5	441	8	US-08-459-455-43	Sequence 43, Appl	185	27	84.4	225	9	US-09-764-868-1102	Sequence 1102, Ap
113	28	87.5	445	16	US-10-767-701-41046	Sequence 41046, A	186	27	84.4	230	15	US-10-276-774-2405	Sequence 2405, Ap
114	28	87.5	446	9	US-09-306-226-30	Sequence 30, Appl	187	27	84.4	235	15	US-10-320-797-3036	Sequence 3036, Ap
115	28	87.5	451	9	US-09-888-243-28	Sequence 28, Appl	188	27	84.4	247	17	US-10-741-600-1591	Sequence 1591, Ap
116	28	87.5	452	14	US-10-205-219-52	Sequence 52, Appl	189	27	84.4	247	17	US-10-741-600-1592	Sequence 1592, Ap
117	28	87.5	456	16	US-10-437-963-124603	Sequence 124603,	190	27	84.4	263	16	US-10-755-889-294	Sequence 294, App
118	28	87.5	457	16	US-10-481-179-8	Sequence 8, Appl	191	27	84.4	280	13	US-10-115-899-5	Sequence 5, Appl
119	28	87.5	466	15	US-10-369-493-2996	Sequence 2996, Ap	192	27	84.4	280	13	US-10-115-899-8	Sequence 8, Appl
120	28	87.5	468	16	US-10-408-765A-1807	Sequence 1807, Ap	193	27	84.4	280	16	US-10-203-532-6	Sequence 6, Appl
121	28	87.5	469	15	US-10-166-349-8	Sequence 8, Appl	194	27	84.4	280	16	US-10-868-270-1	Sequence 1, Appl
122	28	87.5	469	15	US-10-678-986-1	Sequence 1, Appl	195	27	84.4	280	16	US-10-868-270-6	Sequence 6, Appl
123	28	87.5	469	16	US-10-408-765A-1346	Sequence 1346, Ap	196	27	84.4	291	17	US-10-624-727-53	Sequence 53, Appl
124	28	87.5	469	16	US-10-408-765A-2300	Sequence 2300, Ap	197	27	84.4	291	17	US-10-624-727-61	Sequence 61, Appl
125	28	87.5	470	15	US-10-369-493-2942	Sequence 2942, Ap	198	27	84.4	299	9	US-09-815-242-10205	Sequence 10205, A
126	28	87.5	477	14	US-10-120-018-6	Sequence 6, Appl	199	27	84.4	299	15	US-10-282-122A-56595	Sequence 56595, A
127	28	87.5	524	15	US-10-424-599-213918	Sequence 213918,	200	27	84.4	326	16	US-10-425-115-352784	Sequence 352784,
128	28	87.5	530	14	US-10-053-637-12	Sequence 12, Appl	201	27	84.4	341	15	US-10-282-122A-62115	Sequence 62115, A
129	28	87.5	550	16	US-10-425-115-206001	Sequence 206001,	202	27	84.4	342	16	US-10-437-963-185725	Sequence 185725,
130	28	87.5	560	9	US-09-738-626-6264	Sequence 6264, Ap	203	27	84.4	343	15	US-10-282-122A-64623	Sequence 64623, A
131	28	87.5	596	16	US-10-437-963-174585	Sequence 174585,	204	27	84.4	353	16	US-10-778-597-76	Sequence 76, Appl
132	28	87.5	616	15	US-10-108-260A-2508	Sequence 2508, Ap	205	27	84.4	357	15	US-10-282-122A-64960	Sequence 64960, A
133	28	87.5	629	16	US-10-375-010-24	Sequence 24, Appl	206	27	84.4	395	10	US-09-949-427-2	Sequence 2, Appl
134	28	87.5	659	16	US-10-643-795A-141	Sequence 141, App	207	27	84.4	395	10	US-09-949-428-2	Sequence 2, Appl
135	28	87.5	699	17	US-10-482-029-295	Sequence 295, App	208	27	84.4	448	16	US-10-437-963-183864	Sequence 183864,
136	28	87.5	699	17	US-10-948-518-141	Sequence 141, App	209	27	84.4	455	17	US-10-831-070-116	Sequence 116, App
137	28	87.5	722	14	US-10-222-668-4	Sequence 4, Appl	210	27	84.4	489	16	US-10-437-963-104970	Sequence 104970,
138	28	87.5	730	14	US-10-067-632-58	Sequence 58, Appl	211	27	84.4	492	16	US-10-437-963-192093	Sequence 192093,
139	28	87.5	739	16	US-10-437-963-141439	Sequence 141439,	212	27	84.4	496	16	US-10-425-115-304877	Sequence 304877,
140	28	87.5	746	16	US-10-437-963-118707	Sequence 118707,	213	27	84.4	526	16	US-10-481-179-2	Sequence 2, Appl
141	28	87.5	809	16	US-10-489-425-12	Sequence 12, Appl	214	27	84.4	547	16	US-10-437-963-114235	Sequence 114235,
142	28	87.5	822	9	US-09-147-947-6	Sequence 6, Appl	215	27	84.4	552	16	US-10-425-115-216188	Sequence 216188,
143	28	87.5	842	9	US-09-815-242-11950	Sequence 11950, A	216	27	84.4	553	10	US-09-927-239C-1	Sequence 1, Appl
144	28	87.5	842	15	US-10-282-122A-66544	Sequence 66544, A	217	27	84.4	553	17	US-10-703-986-1	Sequence 1, Appl
145	28	87.5	875	16	US-10-757-262-34	Sequence 34, Appl	218	27	84.4	559	14	US-10-163-587A-11	Sequence 11, Appl
146	28	87.5	875	16	US-10-723-860-4046	Sequence 4046, Ap	219	27	84.4	559	14	US-10-205-219-38	Sequence 38, Appl
147	28	87.5	875	17	US-10-843-299-2	Sequence 2, Appl	220	27	84.4	559	14	US-10-369-378-23	Sequence 23, Appl
148	28	87.5	963	15	US-10-464-368-86	Sequence 86, Appl	221	27	84.4	575	15	US-10-104-047-3176	Sequence 3176, Ap
149	28	87.5	963	16	US-10-479-875-5	Sequence 5, Appl	222	27	84.4	595	15	US-10-425-114-63271	Sequence 63271, A
150	28	87.5	1035	16	US-10-437-963-130025	Sequence 130025,	223	27	84.4	602	14	US-10-161-051-143	Sequence 143, App
151	28	87.5	1075	16	US-10-425-115-274886	Sequence 274886,	224	27	84.4	605	9	US-09-782-980-62	Sequence 62, Appl
152	28	87.5	1280	14	US-10-087-887-86	Sequence 86, Appl	225	27	84.4	605	16	US-10-408-765A-632	Sequence 632, App
153	28	87.5	1806	15	US-10-369-493-5622	Sequence 5622, Ap	226	27	84.4	605	16	US-10-706-791-21	Sequence 21, Appl
154	28	87.5	2043	15	US-10-120-801-81	Sequence 81, Appl	227	27	84.4	605	16	US-10-806-018-62	Sequence 62, Appl
155	27	84.4	34	9	US-09-925-297-883	Sequence 883, App	228	27	84.4	658	16	US-10-408-765A-2647	Sequence 2647, Ap
156	27	84.4	50	16	US-10-425-115-367125	Sequence 367125,	229	27	84.4	686	16	US-10-437-963-128149	Sequence 128149,
157	27	84.4	54	17	US-10-926-683-957	Sequence 957, App	230	27	84.4	713	16	US-10-437-963-113424	Sequence 113424,

231	27	84.4	736	15	US-10-250-613-10	Sequence 10, Appl	304	26	81.2	9	10	US-09-791-524-124	Sequence 124, App
232	27	84.4	1011	15	US-10-422-536-111	Sequence 111, Appl	305	26	81.2	9	11	US-09-995-388-47	Sequence 47, Appl
233	27	84.4	1011	15	US-10-232-531-9	Sequence 9, Appl	306	26	81.2	9	13	US-10-013-009-1	Sequence 1, Appl
234	27	84.4	1036	16	US-10-437-963-192091	Sequence 972091,	307	26	81.2	9	13	US-10-033-769-10	Sequence 10, Appl
235	27	84.4	1059	15	US-10-334-143-87	Sequence 87, Appl	308	26	81.2	9	13	US-10-081-258-10	Sequence 10, Appl
236	27	84.4	1137	10	US-09-949-427-405	Sequence 405, App	309	26	81.2	9	13	US-10-080-854-8	Sequence 8, Appl
237	27	84.4	1137	10	US-09-949-428-405	Sequence 405, App	310	26	81.2	9	13	US-10-059-749-56	Sequence 56, Appl
238	27	84.4	2174	14	US-10-087-887-87	Sequence 87, Appl	311	26	81.2	9	13	US-10-038-972A-10	Sequence 10, Appl
239	26	81.2	5	9	US-09-813-484-24	Sequence 24, Appl	312	26	81.2	9	14	US-10-304-160-3	Sequence 3, Appl
240	26	81.2	5	9	US-09-912-609-9	Sequence 9, Appl	313	26	81.2	9	14	US-10-264-374-1	Sequence 1, Appl
241	26	81.2	5	9	US-09-364-597A-37	Sequence 37, Appl	314	26	81.2	9	14	US-10-032-221B-35	Sequence 35, Appl
242	26	81.2	5	14	US-10-046-801-24	Sequence 24, Appl	315	26	81.2	9	14	US-10-375-992-1	Sequence 1, Appl
243	26	81.2	5	14	US-10-304-160-2	Sequence 2, Appl	316	26	81.2	9	15	US-10-269-575-1	Sequence 1, Appl
244	26	81.2	5	17	US-10-838-289-65	Sequence 65, Appl	317	26	81.2	9	15	US-10-269-575-2	Sequence 2, Appl
245	26	81.2	6	9	US-09-912-609-10	Sequence 10, Appl	318	26	81.2	9	15	US-10-269-575-3	Sequence 3, Appl
246	26	81.2	6	9	US-09-912-609-13	Sequence 13, Appl	319	26	81.2	9	15	US-10-269-575-5	Sequence 5, Appl
247	26	81.2	6	9	US-09-364-597A-7	Sequence 7, Appl	320	26	81.2	9	15	US-10-296-879-28	Sequence 28, Appl
248	26	81.2	6	17	US-10-838-289-66	Sequence 66, Appl	321	26	81.2	9	15	US-10-609-217-449	Sequence 449, App
249	26	81.2	6	17	US-10-838-289-67	Sequence 67, Appl	322	26	81.2	9	15	US-10-609-217-450	Sequence 450, App
250	26	81.2	7	9	US-09-364-597A-13	Sequence 13, Appl	323	26	81.2	9	15	US-10-609-217-451	Sequence 451, App
251	26	81.2	7	9	US-09-364-597A-14	Sequence 14, Appl	324	26	81.2	9	15	US-10-609-217-452	Sequence 452, App
252	26	81.2	7	9	US-09-364-597A-30	Sequence 30, Appl	325	26	81.2	9	15	US-10-609-217-1076	Sequence 1076, Ap
253	26	81.2	7	9	US-09-840-277-59	Sequence 59, Appl	326	26	81.2	9	15	US-10-363-208-166	Sequence 166, App
254	26	81.2	7	10	US-09-791-524-148	Sequence 148, App	327	26	81.2	9	15	US-10-652-244-20	Sequence 20, Appl
255	26	81.2	7	14	US-10-264-374-220	Sequence 220, App	328	26	81.2	9	15	US-10-632-388-449	Sequence 449, App
256	26	81.2	7	15	US-10-609-217-444	Sequence 444, App	329	26	81.2	9	15	US-10-632-388-450	Sequence 450, App
257	26	81.2	7	15	US-10-632-388-444	Sequence 444, App	330	26	81.2	9	15	US-10-632-388-451	Sequence 451, App
258	26	81.2	7	15	US-10-651-723-444	Sequence 444, App	331	26	81.2	9	15	US-10-632-388-452	Sequence 452, App
259	26	81.2	7	15	US-10-645-761-444	Sequence 444, App	332	26	81.2	9	15	US-10-632-388-1076	Sequence 1076, Ap
260	26	81.2	7	15	US-10-666-696-444	Sequence 444, App	333	26	81.2	9	15	US-10-668-955-44	Sequence 44, Appl
261	26	81.2	7	15	US-10-653-048-444	Sequence 444, App	334	26	81.2	9	15	US-10-651-723-449	Sequence 449, App
262	26	81.2	7	15	US-10-264-374-220	Sequence 220, App	335	26	81.2	9	15	US-10-651-723-450	Sequence 450, App
263	26	81.2	7	16	US-10-818-246-2	Sequence 2, Appl	336	26	81.2	9	15	US-10-651-723-451	Sequence 451, App
264	26	81.2	7	18	US-10-645-784-444	Sequence 444, App	337	26	81.2	9	15	US-10-651-723-452	Sequence 452, App
265	26	81.2	8	9	US-09-813-484-19	Sequence 19, Appl	338	26	81.2	9	15	US-10-651-723-1076	Sequence 1076, Ap
266	26	81.2	8	9	US-09-813-484-20	Sequence 20, Appl	339	26	81.2	9	15	US-10-645-761-449	Sequence 449, App
267	26	81.2	8	9	US-09-813-484-21	Sequence 21, Appl	340	26	81.2	9	15	US-10-645-761-450	Sequence 450, App
268	26	81.2	8	9	US-09-424-656-5	Sequence 5, Appl	341	26	81.2	9	15	US-10-645-761-451	Sequence 451, App
269	26	81.2	8	9	US-09-364-597A-27	Sequence 27, Appl	342	26	81.2	9	15	US-10-645-761-452	Sequence 452, App
270	26	81.2	8	13	US-10-013-009-4	Sequence 4, Appl	343	26	81.2	9	15	US-10-645-761-1076	Sequence 1076, Ap
271	26	81.2	8	14	US-10-137-435-5	Sequence 5, Appl	344	26	81.2	9	15	US-10-666-696-449	Sequence 449, App
272	26	81.2	8	14	US-10-046-801-19	Sequence 19, Appl	345	26	81.2	9	15	US-10-666-696-450	Sequence 450, App
273	26	81.2	8	14	US-10-046-801-20	Sequence 20, Appl	346	26	81.2	9	15	US-10-666-696-451	Sequence 451, App
274	26	81.2	8	14	US-10-046-801-21	Sequence 21, Appl	347	26	81.2	9	15	US-10-666-696-452	Sequence 452, App
275	26	81.2	8	14	US-10-083-861B-7	Sequence 7, Appl	348	26	81.2	9	15	US-10-666-696-1076	Sequence 1076, Ap
276	26	81.2	8	15	US-10-296-879-5	Sequence 5, Appl	349	26	81.2	9	15	US-10-653-048-449	Sequence 449, App
277	26	81.2	8	15	US-10-296-879-27	Sequence 27, Appl	350	26	81.2	9	15	US-10-653-048-450	Sequence 450, App
278	26	81.2	8	17	US-10-753-729-1	Sequence 1, Appl	351	26	81.2	9	15	US-10-653-048-451	Sequence 451, App
279	26	81.2	9	8	US-08-865-579-56	Sequence 56, Appl	352	26	81.2	9	15	US-10-653-048-452	Sequence 452, App
280	26	81.2	9	9	US-09-746-731-56	Sequence 56, Appl	353	26	81.2	9	15	US-10-653-048-1076	Sequence 1076, Ap
281	26	81.2	9	9	US-09-765-086-1	Sequence 1, Appl	354	26	81.2	9	15	US-10-264-374-1	Sequence 1, Appl
282	26	81.2	9	9	US-09-952-768-44	Sequence 44, Appl	355	26	81.2	9	16	US-10-375-992-1	Sequence 1, Appl
283	26	81.2	9	9	US-09-912-609-31	Sequence 31, Appl	356	26	81.2	9	16	US-10-697-535-1	Sequence 1, Appl
284	26	81.2	9	9	US-09-845-160-5	Sequence 5, Appl	357	26	81.2	9	16	US-10-815-514-33	Sequence 33, Appl
285	26	81.2	9	9	US-09-245-603A-16	Sequence 16, Appl	358	26	81.2	9	16	US-10-729-156-4	Sequence 4, Appl
286	26	81.2	9	9	US-09-364-597A-15	Sequence 15, Appl	359	26	81.2	9	16	US-10-877-930-33	Sequence 33, Appl
287	26	81.2	9	9	US-09-364-597A-16	Sequence 16, Appl	360	26	81.2	9	17	US-10-835-151-1	Sequence 1, Appl
288	26	81.2	9	9	US-09-364-597A-17	Sequence 17, Appl	361	26	81.2	9	17	US-10-835-151-17	Sequence 17, Appl
289	26	81.2	9	9	US-09-364-597A-18	Sequence 18, Appl	362	26	81.2	9	17	US-10-655-755-12	Sequence 12, Appl
290	26	81.2	9	9	US-09-364-597A-33	Sequence 33, Appl	363	26	81.2	9	17	US-10-838-289-87	Sequence 87, Appl
291	26	81.2	9	9	US-09-364-597A-34	Sequence 34, Appl	364	26	81.2	9	17	US-10-838-289-267	Sequence 267, App
292	26	81.2	9	9	US-09-995-388-47	Sequence 47, Appl	365	26	81.2	9	17	US-10-912-764-34	Sequence 34, Appl
293	26	81.2	9	9	US-09-954-697-94	Sequence 94, Appl	366	26	81.2	9	17	US-10-917-709-2	Sequence 2, Appl
294	26	81.2	9	9	US-09-734-628-1	Sequence 1, Appl	367	26	81.2	9	17	US-10-363-205-9	Sequence 9, Appl
295	26	81.2	9	9	US-09-971-798-5	Sequence 5, Appl	368	26	81.2	9	17	US-10-363-205-78	Sequence 78, Appl
296	26	81.2	9	9	US-09-969-192-3	Sequence 3, Appl	369	26	81.2	9	17	US-10-363-205-84	Sequence 84, Appl
297	26	81.2	9	9	US-09-969-192-4	Sequence 4, Appl	370	26	81.2	9	17	US-10-363-205-102	Sequence 102, App
298	26	81.2	9	9	US-09-840-277-14	Sequence 14, Appl	371	26	81.2	9	17	US-10-363-205-220	Sequence 220, App
299	26	81.2	9	9	US-09-840-277-22	Sequence 22, Appl	372	26	81.2	9	17	US-10-363-205-273	Sequence 273, App
300	26	81.2	9	9	US-09-840-277-38	Sequence 38, Appl	373	26	81.2	9	17	US-10-607-595-211	Sequence 211, App
301	26	81.2	9	9	US-09-840-277-62	Sequence 62, Appl	374	26	81.2	9	17	US-10-873-594-33	Sequence 33, Appl
302	26	81.2	9	9	US-09-840-277-63	Sequence 63, Appl	375	26	81.2	9	18	US-10-645-784-449	Sequence 449, App
303	26	81.2	9	10	US-09-801-485-2	Sequence 2, Appl	376	26	81.2	9	18	US-10-645-784-450	Sequence 450, App

377	26	81.2	9	18	US-10-645-784-451	Sequence 451, App	450	15	16	US-10-442-658-11	Sequence 11, Appl
378	26	81.2	9	18	US-10-645-784-452	Sequence 452, App	451	16	10	US-09-801-485-10	Sequence 10, Appl
379	26	81.2	9	18	US-10-645-784-1076	Sequence 1076, App	452	16	15	US-10-353-522A-7	Sequence 7, Appl
380	26	81.2	9	18	US-10-900-399-20	Sequence 20, Appl	453	16	17	US-10-917-709-10	Sequence 10, Appl
381	26	81.2	10	9	US-09-780-142-2	Sequence 2, Appl	454	16	9	US-09-757-908A-16	Sequence 16, Appl
382	26	81.2	10	9	US-09-424-656-6	Sequence 6, Appl	455	18	10	US-09-801-485-8	Sequence 8, Appl
383	26	81.2	10	9	US-09-845-160-14	Sequence 14, Appl	456	18	10	US-09-801-485-9	Sequence 9, Appl
384	26	81.2	10	9	US-09-364-597A-5	Sequence 5, Appl	457	18	15	US-10-430-685-26	Sequence 26, Appl
385	26	81.2	10	9	US-09-870-203A-43	Sequence 43, Appl	458	18	15	US-10-917-709-8	Sequence 8, Appl
386	26	81.2	10	14	US-10-116-067-5	Sequence 5, Appl	459	18	17	US-10-917-709-9	Sequence 9, Appl
387	26	81.2	10	14	US-10-137-435-6	Sequence 6, Appl	460	19	15	US-10-353-522A-3	Sequence 3, Appl
388	26	81.2	10	14	US-10-013-815-21	Sequence 21, Appl	461	19	15	US-10-353-522A-4	Sequence 4, Appl
389	26	81.2	10	14	US-10-429-496-2	Sequence 2, Appl	462	20	9	US-09-205-658-27	Sequence 27, Appl
390	26	81.2	10	14	US-10-429-428-2	Sequence 2, Appl	463	20	10	US-09-844-553A-27	Sequence 27, Appl
391	26	81.2	10	15	US-10-269-575-4	Sequence 4, Appl	464	20	9	US-09-963-693-27	Sequence 27, Appl
392	26	81.2	10	15	US-10-403-337-30	Sequence 30, Appl	465	20	15	US-10-272-437A-10	Sequence 10, Appl
393	26	81.2	10	15	US-10-351-890-30	Sequence 30, Appl	466	20	15	US-10-162-043-2	Sequence 2, Appl
394	26	81.2	10	15	US-10-296-879-6	Sequence 6, Appl	467	21	14	US-10-661-156-525	Sequence 525, App
395	26	81.2	10	15	US-10-296-879-29	Sequence 29, Appl	468	21	17	US-09-864-761-34955	Sequence 34955, A
396	26	81.2	10	15	US-10-296-879-33	Sequence 33, Appl	469	22	15	US-10-353-522A-9	Sequence 9, Appl
397	26	81.2	10	17	US-10-812-238A-21	Sequence 21, Appl	470	22	15	US-10-353-522A-11	Sequence 11, Appl
398	26	81.2	11	9	US-09-765-086-16	Sequence 16, Appl	471	23	9	US-09-969-192-5	Sequence 5, Appl
399	26	81.2	11	9	US-09-424-656-7	Sequence 7, Appl	472	23	14	US-10-162-043-5	Sequence 5, Appl
400	26	81.2	11	9	US-09-424-656-8	Sequence 8, Appl	473	24	8	US-08-902-572-10	Sequence 10, Appl
401	26	81.2	11	9	US-09-424-656-9	Sequence 9, Appl	474	24	9	US-09-862-179A-37	Sequence 37, Appl
402	26	81.2	11	9	US-09-938-623-20	Sequence 20, Appl	475	24	9	US-09-969-192-49	Sequence 49, Appl
403	26	81.2	11	9	US-09-364-597A-10	Sequence 10, Appl	476	24	14	US-10-161-499-8	Sequence 8, Appl
404	26	81.2	11	10	US-09-801-485-1	Sequence 1, Appl	477	24	14	US-10-059-261-281	Sequence 281, App
405	26	81.2	11	10	US-09-791-524-14	Sequence 14, Appl	478	24	14	US-10-059-261-283	Sequence 283, App
406	26	81.2	11	14	US-10-137-435-7	Sequence 7, Appl	479	24	14	US-10-059-261-301	Sequence 301, App
407	26	81.2	11	14	US-10-137-435-8	Sequence 8, Appl	480	24	14	US-10-059-261-302	Sequence 302, App
408	26	81.2	11	14	US-10-137-435-9	Sequence 9, Appl	481	24	15	US-10-378-173-9	Sequence 9, Appl
409	26	81.2	11	14	US-10-264-374-16	Sequence 16, Appl	482	24	15	US-10-275-427A-42	Sequence 42, App
410	26	81.2	11	14	US-10-375-992-16	Sequence 16, Appl	483	24	16	US-10-627-649-281	Sequence 281, App
411	26	81.2	11	14	US-10-136-819-9	Sequence 9, Appl	484	24	16	US-10-627-649-283	Sequence 283, App
412	26	81.2	11	15	US-10-136-187-10	Sequence 10, Appl	485	24	16	US-10-627-649-301	Sequence 301, App
413	26	81.2	11	15	US-10-136-187-14	Sequence 14, Appl	486	24	16	US-10-627-649-302	Sequence 302, App
414	26	81.2	11	15	US-10-296-879-7	Sequence 7, Appl	487	24	17	US-10-744-964-8	Sequence 8, Appl
415	26	81.2	11	15	US-10-296-879-8	Sequence 8, Appl	488	25	14	US-10-059-261-311	Sequence 311, App
416	26	81.2	11	15	US-10-296-879-9	Sequence 9, Appl	489	25	14	US-10-059-261-312	Sequence 312, App
417	26	81.2	11	15	US-10-264-374-16	Sequence 16, Appl	490	25	14	US-10-059-261-313	Sequence 313, App
418	26	81.2	11	16	US-10-375-992-16	Sequence 16, Appl	491	25	14	US-10-059-261-314	Sequence 314, App
419	26	81.2	11	16	US-10-469-131-27	Sequence 27, Appl	492	25	15	US-10-353-522A-1	Sequence 1, Appl
420	26	81.2	11	17	US-10-792-390-2	Sequence 2, Appl	493	25	16	US-10-627-649-311	Sequence 311, App
421	26	81.2	11	17	US-10-850-873-10	Sequence 10, Appl	494	25	16	US-10-627-649-312	Sequence 312, App
422	26	81.2	11	17	US-10-850-873-14	Sequence 14, Appl	495	25	16	US-10-627-649-313	Sequence 313, App
423	26	81.2	11	17	US-10-912-764-35	Sequence 35, Appl	496	25	16	US-10-627-649-314	Sequence 314, App
424	26	81.2	11	17	US-10-917-709-1	Sequence 1, Appl	497	26	9	US-09-424-656-3	Sequence 3, Appl
425	26	81.2	12	9	US-09-424-656-10	Sequence 10, Appl	498	26	14	US-10-137-435-3	Sequence 3, Appl
426	26	81.2	12	9	US-09-969-192-79	Sequence 79, Appl	499	26	15	US-10-296-879-3	Sequence 3, Appl
427	26	81.2	12	14	US-10-137-435-10	Sequence 10, Appl	500	26	15	US-09-424-656-2	Sequence 2, Appl
428	26	81.2	12	14	US-10-059-261-282	Sequence 282, App	501	26	9	US-09-424-656-4	Sequence 4, Appl
429	26	81.2	12	14	US-10-059-261-315	Sequence 315, App	502	26	10	US-09-984-276-253	Sequence 253, App
430	26	81.2	12	14	US-10-059-261-316	Sequence 316, App	503	26	10	US-09-984-271-253	Sequence 253, App
431	26	81.2	12	15	US-10-296-879-10	Sequence 10, Appl	504	26	14	US-10-137-435-2	Sequence 2, Appl
432	26	81.2	12	15	US-10-296-879-15	Sequence 15, Appl	505	26	14	US-10-137-435-4	Sequence 4, Appl
433	26	81.2	12	16	US-10-627-649-282	Sequence 282, App	506	26	14	US-10-216-556-15	Sequence 15, Appl
434	26	81.2	12	16	US-10-627-649-315	Sequence 315, App	507	26	15	US-10-296-879-2	Sequence 2, Appl
435	26	81.2	12	16	US-10-627-649-316	Sequence 316, App	508	26	15	US-10-296-879-4	Sequence 4, Appl
436	26	81.2	13	9	US-09-949-474-16	Sequence 16, Appl	509	26	15	US-10-296-879-20	Sequence 20, Appl
437	26	81.2	13	9	US-09-949-474-17	Sequence 17, Appl	510	26	15	US-10-296-879-22	Sequence 22, Appl
438	26	81.2	13	10	US-09-791-524-13	Sequence 13, Appl	511	26	15	US-10-296-879-24	Sequence 24, Appl
439	26	81.2	13	10	US-09-791-524-23	Sequence 23, Appl	512	26	17	US-10-486-480-15	Sequence 15, Appl
440	26	81.2	13	17	US-10-638-289-226	Sequence 226, App	513	26	9	US-09-424-656-1	Sequence 1, Appl
441	26	81.2	13	17	US-10-607-595-118	Sequence 118, App	514	26	14	US-10-137-435-1	Sequence 1, Appl
442	26	81.2	13	18	US-10-948-707-807	Sequence 807, App	515	26	15	US-10-296-879-1	Sequence 1, Appl
443	26	81.2	14	9	US-09-969-192-68	Sequence 68, Appl	516	26	15	US-10-296-879-23	Sequence 23, Appl
444	26	81.2	14	10	US-09-999-724-93	Sequence 93, Appl	517	26	9	US-09-993-844-17	Sequence 17, Appl
445	26	81.2	14	13	US-10-038-972A-1	Sequence 1, Appl	518	26	14	US-10-121-282-39	Sequence 39, Appl
446	26	81.2	14	15	US-10-181-937-15	Sequence 15, Appl	519	26	15	US-10-633-438-3	Sequence 3, Appl
447	26	81.2	15	9	US-09-912-609-130	Sequence 130, App	520	26	17	US-10-693-164-5	Sequence 5, Appl
448	26	81.2	15	9	US-09-969-192-31	Sequence 31, Appl	521	26	18	US-10-901-772-3	Sequence 3, Appl
449	26	81.2	15	10	US-09-791-524-22	Sequence 22, Appl	522	26	19	US-11-026-435-17	Sequence 17, Appl

523	26	81.2	30	15	US-10-424-599-164382	Sequence 164382,	596	74	15	US-10-424-599-255452	Sequence 255452,
524	26	81.2	32	16	US-10-767-701-59248	Sequence 59248, A	597	74	16	US-10-437-963-173630	Sequence 173630,
525	26	81.2	34	15	US-10-424-599-241038	Sequence 241038,	598	75	14	US-10-156-761-7656	Sequence 7656, Ap
526	26	81.2	35	9	US-09-205-658-30	Sequence 30, Appl	599	75	15	US-10-424-599-269022	Sequence 269022,
527	26	81.2	35	9	US-09-844-353A-30	Sequence 30, Appl	600	75	16	US-10-425-115-268458	Sequence 268458,
528	26	81.2	35	10	US-09-963-693-30	Sequence 30, Appl	601	75	16	US-10-425-115-318005	Sequence 318005,
529	26	81.2	35	15	US-10-424-599-273942	Sequence 273942,	602	76	16	US-10-437-963-148522	Sequence 148522,
530	26	81.2	37	16	US-10-425-115-195099	Sequence 195099,	603	76	16	US-10-425-115-368412	Sequence 368412,
531	26	81.2	39	9	US-09-790-264-63	Sequence 63, Appl	604	77	11	US-09-855-604-326	Sequence 326, App
532	26	81.2	39	9	US-09-757-908A-3	Sequence 3, Appli	605	77	12	US-09-855-604-326	Sequence 326, App
533	26	81.2	39	14	US-10-029-386-28855	Sequence 28855, A	606	77	15	US-10-335-977-6463	Sequence 6463, Ap
534	26	81.2	39	14	US-10-029-353-63	Sequence 63, Appl	607	77	16	US-10-425-115-263112	Sequence 263112,
535	26	81.2	39	17	US-10-900-926-63	Sequence 63, Appl	608	78	16	US-10-425-115-206141	Sequence 206141,
536	26	81.2	40	16	US-10-425-115-361595	Sequence 361595,	609	78	16	US-10-425-115-280058	Sequence 280058,
537	26	81.2	41	16	US-10-425-115-362424	Sequence 362424,	610	79	15	US-10-424-599-266713	Sequence 266713,
538	26	81.2	41	16	US-10-425-115-367745	Sequence 367745,	611	79	16	US-10-425-115-242278	Sequence 242278,
539	26	81.2	43	9	US-09-205-658-28	Sequence 28, Appl	612	79	16	US-10-425-115-281115	Sequence 281115,
540	26	81.2	43	9	US-09-844-353A-28	Sequence 28, Appl	613	80	15	US-10-424-599-193396	Sequence 193396,
541	26	81.2	43	10	US-09-963-693-28	Sequence 28, Appl	614	80	15	US-10-424-599-239157	Sequence 239157,
542	26	81.2	46	8	US-08-855-579-13	Sequence 13, Appl	615	81	15	US-10-424-599-248667	Sequence 248667,
543	26	81.2	46	9	US-09-746-731-13	Sequence 13, Appl	616	81	15	US-10-424-599-173664	Sequence 173664,
544	26	81.2	46	13	US-10-059-749-13	Sequence 13, Appl	617	83	16	US-10-437-963-114667	Sequence 114667,
545	26	81.2	46	15	US-10-261-161-34	Sequence 34, Appl	618	83	16	US-10-425-115-267701	Sequence 267701,
546	26	81.2	46	16	US-10-428-817A-215	Sequence 215, App	619	83	16	US-10-425-115-358490	Sequence 358490,
547	26	81.2	46	18	US-10-694-243-10	Sequence 10, Appl	620	84	16	US-10-425-115-205494	Sequence 205494,
548	26	81.2	48	15	US-10-424-599-161655	Sequence 161655,	621	86	16	US-10-767-701-33462	Sequence 33462, A
549	26	81.2	49	16	US-10-767-701-50090	Sequence 50090, A	622	86	14	US-10-029-386-31442	Sequence 31442, A
550	26	81.2	50	9	US-09-796-692-747	Sequence 747, App	623	86	16	US-10-767-701-50776	Sequence 50776, A
551	26	81.2	50	14	US-10-040-862-747	Sequence 747, App	624	88	15	US-10-074-024-361	Sequence 361, App
552	26	81.2	50	15	US-10-057-475B-747	Sequence 747, App	625	88	16	US-10-437-963-190260	Sequence 190260,
553	26	81.2	50	15	US-10-154-884B-747	Sequence 747, App	626	88	16	US-10-425-115-34644	Sequence 34644,
554	26	81.2	50	16	US-10-764-324-747	Sequence 747, App	627	89	14	US-10-101-464A-798	Sequence 798, App
555	26	81.2	51	16	US-10-425-115-309902	Sequence 309902,	628	89	16	US-10-425-115-281794	Sequence 281794,
556	26	81.2	53	9	US-09-790-364-60	Sequence 60, Appl	629	89	17	US-10-864-252-798	Sequence 798, App
557	26	81.2	53	14	US-10-269-353-60	Sequence 60, Appl	630	90	9	US-09-925-297-676	Sequence 676, App
558	26	81.2	53	15	US-10-424-599-266660	Sequence 266660,	631	90	15	US-10-424-599-158447	Sequence 158447,
559	26	81.2	53	17	US-10-900-926-60	Sequence 60, Appl	632	91	16	US-10-437-963-157284	Sequence 157284,
560	26	81.2	55	16	US-10-425-115-324625	Sequence 324625,	633	91	16	US-10-425-115-333740	Sequence 333740,
561	26	81.2	56	15	US-10-424-599-274959	Sequence 274959,	634	92	16	US-10-425-115-331141	Sequence 331141,
562	26	81.2	57	14	US-10-029-386-28354	Sequence 28354, A	635	92	16	US-10-425-115-331141	Sequence 177520,
563	26	81.2	57	15	US-10-424-599-254444	Sequence 254444,	636	93	15	US-10-462-691-24	Sequence 24, Appl
564	26	81.2	57	16	US-10-425-115-249920	Sequence 249920,	637	94	15	US-10-289-763-1290	Sequence 1290, Ap
565	26	81.2	57	18	US-10-969-719A-79	Sequence 79, Appl	638	94	15	US-10-424-599-216267	Sequence 216267,
566	26	81.2	58	9	US-09-864-761-41764	Sequence 41764, A	639	94	15	US-10-424-599-279209	Sequence 279209,
567	26	81.2	58	11	US-09-864-408A-6776	Sequence 6776, Ap	640	94	16	US-10-437-963-126730	Sequence 126730,
568	26	81.2	58	15	US-10-424-599-284915	Sequence 284915,	641	94	16	US-10-425-115-257240	Sequence 257240,
569	26	81.2	59	16	US-10-425-115-350773	Sequence 350773,	642	94	16	US-10-425-115-343409	Sequence 343409,
570	26	81.2	60	15	US-10-424-599-264640	Sequence 264640,	643	95	15	US-10-424-599-212527	Sequence 212527,
571	26	81.2	61	16	US-10-437-963-125047	Sequence 125047,	644	95	15	US-10-335-977-6467	Sequence 6467, Ap
572	26	81.2	62	16	US-10-425-115-270384	Sequence 270384,	645	96	16	US-10-437-963-201926	Sequence 201926,
573	26	81.2	63	9	US-09-790-264-44	Sequence 44, Appl	646	96	16	US-10-437-963-204226	Sequence 204226,
574	26	81.2	63	9	US-09-864-761-42848	Sequence 42848, A	647	97	11	US-09-855-604-369	Sequence 369, App
575	26	81.2	63	9	US-09-757-908A-2	Sequence 2, Appli	648	97	12	US-09-855-604-369	Sequence 369, App
576	26	81.2	63	14	US-10-269-353-44	Sequence 44, Appl	649	97	16	US-10-425-115-207049	Sequence 207049,
577	26	81.2	63	15	US-10-424-599-156271	Sequence 156271,	650	97	16	US-10-425-115-237373	Sequence 237373,
578	26	81.2	63	16	US-10-425-115-219640	Sequence 219640,	651	97	16	US-10-425-115-250608	Sequence 250608,
579	26	81.2	63	17	US-10-900-926-44	Sequence 44, Appl	652	97	16	US-10-425-115-361081	Sequence 361081,
580	26	81.2	65	9	US-09-925-299-1336	Sequence 1336, Ap	653	98	16	US-10-437-963-131411	Sequence 131411,
581	26	81.2	65	10	US-09-925-299-1336	Sequence 1336, Ap	654	98	16	US-10-437-963-194385	Sequence 194385,
582	26	81.2	65	11	US-09-864-408A-8516	Sequence 8516, Ap	655	99	11	US-09-855-604-324	Sequence 324, App
583	26	81.2	65	14	US-10-029-386-30177	Sequence 30177, A	656	99	12	US-09-855-604-324	Sequence 324, App
584	26	81.2	65	15	US-10-355-161A-12	Sequence 12, Appl	657	99	16	US-10-437-963-196555	Sequence 196555,
585	26	81.2	65	16	US-10-425-115-294932	Sequence 294932,	658	100	16	US-10-437-963-153969	Sequence 153969,
586	26	81.2	66	16	US-10-425-115-208756	Sequence 208756,	659	102	16	US-10-425-115-194836	Sequence 194836,
587	26	81.2	67	16	US-10-767-701-35906	Sequence 35906, A	660	102	16	US-10-425-115-231988	Sequence 231988,
588	26	81.2	68	15	US-10-424-599-195952	Sequence 195952,	661	103	15	US-10-104-047-2614	Sequence 2614, Ap
589	26	81.2	68	16	US-10-425-115-287469	Sequence 287469,	662	103	15	US-10-424-599-146895	Sequence 146895,
590	26	81.2	72	10	US-09-764-891-3178	Sequence 3178, Ap	663	103	16	US-10-437-963-191754	Sequence 191754,
591	26	81.2	72	16	US-10-437-963-157372	Sequence 157372,	664	103	16	US-10-767-701-33279	Sequence 33279, A
592	26	81.2	72	16	US-10-425-115-202358	Sequence 202358,	665	103	16	US-10-767-701-54979	Sequence 54979, A
593	26	81.2	72	16	US-10-425-115-293042	Sequence 293042,	666	103	16	US-10-425-115-203484	Sequence 203484,
594	26	81.2	72	18	US-10-969-719A-189	Sequence 189, App	667	103	16	US-10-425-115-21845	Sequence 21845,
595	26	81.2	73	16	US-10-425-115-266720	Sequence 266720,	668	104	15	US-10-425-114-42364	Sequence 42364, A

669	26	81.2	104	18	US-10-756-149-5562	Sequence 5562, Ap	742	149	15	US-10-424-599-187566	Sequence 187566,
670	26	81.2	106	16	US-10-425-115-344198	Sequence 344198,	743	150	15	US-10-425-114-51314	Sequence 51314, A
671	26	81.2	107	15	US-10-424-599-150377	Sequence 150377,	744	151	16	US-10-437-963-195875	Sequence 195875,
672	26	81.2	107	16	US-10-767-701-57415	Sequence 57415, A	745	150	16	US-10-767-701-36597	Sequence 36597, A
673	26	81.2	108	9	US-09-925-999-792	Sequence 792, App	746	150	16	US-10-425-115-214955	Sequence 214955,
674	26	81.2	108	16	US-09-925-299-792	Sequence 792, App	747	150	16	US-10-425-115-283827	Sequence 283827,
675	26	81.2	108	16	US-10-800-197-113	Sequence 113, App	748	151	15	US-10-424-599-183242	Sequence 183242,
676	26	81.2	108	16	US-10-800-197-129	Sequence 129, App	749	151	15	US-10-425-114-49863	Sequence 48863, A
677	26	81.2	108	16	US-10-425-115-365069	Sequence 365069,	750	151	15	US-10-425-114-49863	Sequence 48863, A
678	26	81.2	108	20	US-11-021-438-15	Sequence 15, Appl	751	152	10	US-09-764-891-4604	Sequence 4604, Ap
679	26	81.2	109	16	US-10-437-963-202699	Sequence 202699,	752	152	16	US-10-425-115-223763	Sequence 223763,
680	26	81.2	110	16	US-10-437-963-114763	Sequence 114763,	753	153	16	US-10-437-963-181162	Sequence 181162,
681	26	81.2	110	16	US-10-800-197-125	Sequence 125, App	754	154	15	US-10-424-599-219400	Sequence 219400,
682	26	81.2	111	16	US-10-425-115-203244	Sequence 203244,	755	154	16	US-10-767-701-57067	Sequence 57067, A
683	26	81.2	113	15	US-10-424-599-241435	Sequence 241435,	756	155	16	US-10-425-115-298221	Sequence 298221,
684	26	81.2	114	15	US-10-424-599-221678	Sequence 221678,	757	156	15	US-10-425-114-71810	Sequence 71810, A
685	26	81.2	115	16	US-10-425-115-362725	Sequence 362725,	758	157	15	US-10-424-599-202312	Sequence 202312,
686	26	81.2	115	18	US-10-916-758-34	Sequence 34, Appl	759	158	9	US-09-834-759-508	Sequence 508, App
687	26	81.2	115	18	US-10-916-758-62	Sequence 62, Appl	760	158	13	US-10-007-805-508	Sequence 508, App
688	26	81.2	116	16	US-10-425-115-249415	Sequence 249415,	761	158	14	US-10-076-622-508	Sequence 508, App
689	26	81.2	117	16	US-10-767-701-42273	Sequence 42273, A	762	158	14	US-10-124-805-508	Sequence 508, App
690	26	81.2	118	11	US-09-833-245-1578	Sequence 1578, Ap	763	158	15	US-10-424-599-281989	Sequence 281989,
691	26	81.2	119	15	US-10-936-115-945	Sequence 945, App	764	159	15	US-10-425-114-64554	Sequence 64554, A
692	26	81.2	119	16	US-10-437-963-131983	Sequence 131983,	765	159	16	US-10-437-963-106361	Sequence 106361,
693	26	81.2	119	16	US-10-425-115-267015	Sequence 267015,	766	160	16	US-10-437-963-127657	Sequence 127657,
694	26	81.2	120	15	US-10-424-599-169422	Sequence 169422,	767	160	10	US-09-372-348-23	Sequence 23, Appl
695	26	81.2	121	15	US-10-424-599-168636	Sequence 168636,	768	160	13	US-10-000-066-7	Sequence 7, Appl
696	26	81.2	121	16	US-10-767-701-49824	Sequence 49824, A	769	160	16	US-10-723-860-522	Sequence 522, App
697	26	81.2	122	16	US-10-437-963-127558	Sequence 127558,	770	160	17	US-10-902-777-7	Sequence 7, Appl
698	26	81.2	123	16	US-10-767-701-52152	Sequence 52152, A	771	161	16	US-10-425-115-259299	Sequence 259299,
699	26	81.2	123	16	US-10-425-115-239470	Sequence 239470,	772	163	14	US-10-101-464A-637	Sequence 637, App
700	26	81.2	125	15	US-10-424-599-144909	Sequence 144909,	773	163	15	US-10-424-599-210579	Sequence 210579,
701	26	81.2	125	16	US-10-425-115-298225	Sequence 298225,	774	163	17	US-10-864-253-637	Sequence 637, App
702	26	81.2	125	16	US-10-425-115-337679	Sequence 337679,	775	164	16	US-10-437-963-196768	Sequence 196768,
703	26	81.2	126	8	US-08-849-303-27	Sequence 27, Appl	776	164	16	US-10-767-701-56918	Sequence 56918, A
704	26	81.2	126	15	US-10-424-599-154821	Sequence 154821,	777	166	15	US-10-424-599-277844	Sequence 277844,
705	26	81.2	126	16	US-10-437-963-176801	Sequence 176801,	778	167	16	US-10-437-963-156378	Sequence 156378,
706	26	81.2	126	16	US-10-655-136-27	Sequence 27, Appl	779	167	16	US-10-767-701-32577	Sequence 32577, A
707	26	81.2	127	16	US-10-437-963-139052	Sequence 139052,	780	169	16	US-10-425-115-230967	Sequence 230967,
708	26	81.2	127	16	US-10-425-115-201997	Sequence 201997,	781	170	15	US-10-264-049-4186	Sequence 4186, Ap
709	26	81.2	127	16	US-10-425-115-250609	Sequence 250609,	782	175	13	US-10-087-192-1101	Sequence 1101, Ap
710	26	81.2	128	11	US-09-855-604-382	Sequence 382, App	783	175	15	US-10-377-079-66	Sequence 66, Appl
711	26	81.2	128	12	US-09-855-604-382	Sequence 382, App	784	175	16	US-10-437-963-108928	Sequence 108928,
712	26	81.2	128	15	US-10-094-749-2788	Sequence 2788, Ap	785	176	16	US-10-425-115-367391	Sequence 367391,
713	26	81.2	128	16	US-10-425-115-237010	Sequence 237010,	786	177	15	US-10-335-977-7815	Sequence 7815, Ap
714	26	81.2	129	15	US-10-424-599-276829	Sequence 276829,	787	179	11	US-09-833-245-1853	Sequence 1853, Ap
715	26	81.2	129	16	US-10-767-701-31819	Sequence 31819, A	788	179	15	US-10-425-114-53545	Sequence 53545, A
716	26	81.2	130	16	US-10-425-115-223454	Sequence 223454,	789	179	16	US-10-437-963-109901	Sequence 109901,
717	26	81.2	131	16	US-10-437-963-131679	Sequence 131679,	790	179	16	US-10-767-701-39934	Sequence 39934, A
718	26	81.2	131	16	US-10-425-115-275326	Sequence 275326,	791	180	15	US-10-424-599-153366	Sequence 153366,
719	26	81.2	131	16	US-10-739-930-5978	Sequence 5978, Ap	792	181	15	US-10-424-599-219956	Sequence 219956,
720	26	81.2	132	16	US-10-437-963-131932	Sequence 131932,	793	181	16	US-10-437-963-119756	Sequence 119756,
721	26	81.2	132	16	US-10-437-963-182179	Sequence 182179,	794	181	16	US-10-767-701-41045	Sequence 41045, A
722	26	81.2	134	16	US-10-425-115-221918	Sequence 221918,	795	182	16	US-10-437-963-169716	Sequence 169716,
723	26	81.2	134	16	US-10-425-115-302820	Sequence 302820,	796	182	16	US-10-425-115-201775	Sequence 201775,
724	26	81.2	135	15	US-10-276-774-1436	Sequence 1436, Ap	797	184	15	US-10-424-599-150334	Sequence 150334,
725	26	81.2	137	16	US-10-767-701-55617	Sequence 55617, A	798	184	16	US-10-437-963-116550	Sequence 116550,
726	26	81.2	137	16	US-10-425-115-322625	Sequence 322625,	799	185	16	US-10-437-963-168735	Sequence 168735,
727	26	81.2	137	17	US-10-732-923-11042	Sequence 11042, A	800	186	15	US-10-424-599-287100	Sequence 287100,
728	26	81.2	138	16	US-10-437-963-144054	Sequence 144054,	801	186	16	US-10-424-599-153956	Sequence 153956, A
729	26	81.2	138	16	US-10-425-115-208769	Sequence 208769,	802	187	16	US-10-767-701-55996	Sequence 55996, A
730	26	81.2	139	15	US-10-424-599-279945	Sequence 279945,	803	188	16	US-10-437-963-136910	Sequence 136910,
731	26	81.2	140	16	US-10-437-963-169955	Sequence 169955,	804	189	15	US-10-424-599-233126	Sequence 233126,
732	26	81.2	141	15	US-10-424-599-179258	Sequence 179258,	805	189	16	US-10-437-963-185939	Sequence 185939,
733	26	81.2	141	16	US-10-425-115-265061	Sequence 265061,	806	190	16	US-10-767-701-59779	Sequence 59779, A
734	26	81.2	142	16	US-10-767-701-40951	Sequence 40951, A	807	191	10	US-09-984-276-161	Sequence 161, App
735	26	81.2	143	16	US-10-425-115-336583	Sequence 336583,	808	191	10	US-09-833-245-1543	Sequence 1543, Ap
736	26	81.2	144	16	US-10-437-963-156147	Sequence 156147,	809	191	16	US-10-437-963-186301	Sequence 186301,
737	26	81.2	144	16	US-10-425-115-250613	Sequence 250613,	810	191	16	US-10-767-701-43126	Sequence 43126, A
738	26	81.2	144	16	US-10-425-115-265663	Sequence 265663,	811	191	16	US-10-425-115-340441	Sequence 340441,
739	26	81.2	145	16	US-10-425-115-365185	Sequence 365185,	812	192	15	US-10-624-380-10	Sequence 10, Appl
740	26	81.2	147	16	US-10-437-963-173888	Sequence 173888,	813	192	15	US-10-624-380-15	Sequence 15, Appl
741	26	81.2	148	15	US-10-425-114-36645	Sequence 36645, A	814	192	16	US-10-767-701-40878	Sequence 40878, A

815	26	81.2	192	16	US-10-425-115-203688	Sequence 203688,	888	26	81.2	235	15	US-10-412-699B-320	Sequence 320, App
816	26	81.2	192	16	US-10-425-115-364493	Sequence 364493,	889	26	81.2	235	16	US-10-425-115-258496	Sequence 258496,
817	26	81.2	193	16	US-10-425-115-281192	Sequence 281192,	890	26	81.2	235	16	US-10-425-115-322434	Sequence 322434,
818	26	81.2	193	16	US-10-425-115-325225	Sequence 325225,	891	26	81.2	235	16	US-10-425-115-331505	Sequence 331505,
819	26	81.2	194	15	US-10-424-599-142883	Sequence 142883,	892	26	81.2	236	16	US-10-437-963-192857	Sequence 192857,
820	26	81.2	194	15	US-10-424-599-208763	Sequence 208763,	893	26	81.2	236	16	US-10-425-115-268384	Sequence 268384,
821	26	81.2	198	16	US-10-437-963-156432	Sequence 156432,	894	26	81.2	236	16	US-10-425-115-314872	Sequence 314872,
822	26	81.2	198	16	US-10-437-963-202039	Sequence 202039,	895	26	81.2	237	15	US-10-291-172-610	Sequence 610, App
823	26	81.2	198	16	US-10-767-701-36449	Sequence 36449, A	896	26	81.2	237	15	US-10-221-278-610	Sequence 610, App
824	26	81.2	199	9	US-09-948-138-2	Sequence 2, Appli	897	26	81.2	237	16	US-10-437-963-174095	Sequence 174095,
825	26	81.2	200	9	US-09-948-138-2	Sequence 1361, A	898	26	81.2	237	16	US-10-437-963-174095	Sequence 174095,
826	26	81.2	200	15	US-10-369-493-13361	Sequence 13361, A	899	26	81.2	238	16	US-10-767-701-36961	Sequence 36961, A
827	26	81.2	200	16	US-10-425-115-322432	Sequence 322432,	900	26	81.2	238	16	US-10-437-963-117054	Sequence 117054,
828	26	81.2	201	16	US-10-767-701-62667	Sequence 62667, A	901	26	81.2	238	17	US-10-425-115-220593	Sequence 220593,
829	26	81.2	201	16	US-10-425-115-250606	Sequence 250606,	902	26	81.2	239	16	US-10-890-629-14	Sequence 14, Appl
830	26	81.2	202	16	US-10-437-963-204166	Sequence 204166,	903	26	81.2	239	20	US-10-425-115-327249	Sequence 327249,
831	26	81.2	203	11	US-09-855-604-334	Sequence 334, App	904	26	81.2	239	20	US-11-021-438-29	Sequence 29, Appl
832	26	81.2	203	11	US-09-855-604-334	Sequence 352, App	905	26	81.2	241	16	US-10-437-963-114896	Sequence 114896,
833	26	81.2	203	12	US-09-855-604-334	Sequence 334, App	906	26	81.2	241	16	US-10-437-963-117532	Sequence 117532,
834	26	81.2	203	12	US-09-855-604-352	Sequence 352, App	907	26	81.2	242	20	US-09-880-748-1742	Sequence 1742, Ap
835	26	81.2	203	15	US-10-425-114-38712	Sequence 38712, A	908	26	81.2	242	14	US-10-112-221A-6	Sequence 6, Appli
836	26	81.2	205	10	US-09-903-190-122	Sequence 192, App	909	26	81.2	242	15	US-10-293-418-1742	Sequence 1742, Ap
837	26	81.2	205	11	US-09-978-360A-778	Sequence 778, App	910	26	81.2	242	15	US-10-156-214A-18	Sequence 18, Appl
838	26	81.2	205	17	US-10-930-331-122	Sequence 122, App	911	26	81.2	242	15	US-10-424-599-174184	Sequence 174184,
839	26	81.2	209	15	US-10-225-066A-476	Sequence 476, App	912	26	81.2	242	16	US-10-425-114-38011	Sequence 38011, A
840	26	81.2	209	15	US-10-421-138A-99	Sequence 99, Appl	913	26	81.2	242	9	US-10-437-963-184237	Sequence 184237,
841	26	81.2	209	15	US-10-374-780A-2184	Sequence 2184, Ap	914	26	81.2	243	10	US-09-834-759-507	Sequence 507, App
842	26	81.2	209	15	US-10-412-699B-1922	Sequence 1922, Ap	915	26	81.2	243	13	US-10-007-805-507	Sequence 507, App
843	26	81.2	209	16	US-10-437-963-147167	Sequence 147167,	916	26	81.2	243	14	US-10-076-622-507	Sequence 507, App
844	26	81.2	209	17	US-10-230-415-99	Sequence 99, App	917	26	81.2	243	14	US-10-124-805-507	Sequence 507, App
845	26	81.2	209	18	US-10-225-066A-476	Sequence 476, App	918	26	81.2	243	14	US-10-148-671-22	Sequence 22, Appl
846	26	81.2	210	15	US-10-425-114-39015	Sequence 39015, A	919	26	81.2	243	15	US-10-293-418-1863	Sequence 1863, Ap
847	26	81.2	210	16	US-10-767-701-34322	Sequence 34322, A	920	26	81.2	246	15	US-10-425-114-44956	Sequence 44956, A
848	26	81.2	211	15	US-10-425-114-54759	Sequence 54759, A	921	26	81.2	243	15	US-10-424-599-199510	Sequence 199510,
849	26	81.2	212	16	US-10-425-115-260569	Sequence 260569,	922	26	81.2	244	15	US-10-820-474A-23	Sequence 23, Appl
850	26	81.2	213	14	US-10-156-761-10641	Sequence 10641, A	923	26	81.2	244	15	US-10-424-599-166913	Sequence 166913,
851	26	81.2	213	15	US-10-425-114-50313	Sequence 50313, A	924	26	81.2	245	16	US-10-424-599-174186	Sequence 174186,
852	26	81.2	214	16	US-10-437-963-126945	Sequence 126945,	925	26	81.2	246	15	US-10-800-197-11	Sequence 11, Appl
853	26	81.2	214	16	US-10-425-115-336322	Sequence 336322,	926	26	81.2	246	16	US-10-425-114-44956	Sequence 44956, A
854	26	81.2	215	9	US-09-764-587A-4	Sequence 4, Appli	927	26	81.2	247	15	US-10-437-963-117373	Sequence 117373,
855	26	81.2	215	15	US-10-424-599-247945	Sequence 247945,	928	26	81.2	247	15	US-10-282-122A-67885	Sequence 67885, A
856	26	81.2	215	15	US-10-424-599-273845	Sequence 273845,	929	26	81.2	247	16	US-10-425-114-65753	Sequence 65753, A
857	26	81.2	215	16	US-10-437-963-181643	Sequence 181643,	930	26	81.2	247	16	US-10-767-701-43492	Sequence 43492, A
858	26	81.2	216	10	US-09-935-703-10	Sequence 10, Appl	931	26	81.2	248	15	US-10-425-115-279555	Sequence 279555,
859	26	81.2	216	10	US-10-424-599-240748	Sequence 240748,	932	26	81.2	248	16	US-10-425-114-38448	Sequence 38448, A
860	26	81.2	216	16	US-10-425-115-285678	Sequence 285678,	933	26	81.2	249	10	US-10-425-115-276941	Sequence 276941,
861	26	81.2	217	9	US-09-755-016-2	Sequence 2, Appli	934	26	81.2	249	10	US-09-880-748-723	Sequence 723, App
862	26	81.2	217	20	US-11-049-613-2	Sequence 2, Appli	935	26	81.2	249	15	US-09-880-748-1215	Sequence 1215, Ap
863	26	81.2	218	16	US-10-437-963-103015	Sequence 103015,	936	26	81.2	249	15	US-10-293-418-723	Sequence 723, App
864	26	81.2	218	16	US-10-425-115-314618	Sequence 314618,	937	26	81.2	249	16	US-10-293-418-1215	Sequence 1215, Ap
865	26	81.2	219	16	US-10-263-828-94	Sequence 94, Appl	938	26	81.2	250	17	US-10-437-963-167520	Sequence 167520,
866	26	81.2	220	14	US-10-425-114-69245	Sequence 69245, A	939	26	81.2	250	17	US-10-935-290-117	Sequence 117, App
867	26	81.2	221	15	US-10-425-114-69245	Sequence 69245, A	940	26	81.2	251	10	US-09-880-748-1207	Sequence 1207, Ap
868	26	81.2	221	16	US-10-767-701-42431	Sequence 42431, A	941	26	81.2	251	15	US-10-293-418-1207	Sequence 1207, Ap
869	26	81.2	222	15	US-10-437-963-151389	Sequence 151389,	942	26	81.2	251	16	US-10-800-197-3	Sequence 3, Appli
870	26	81.2	223	15	US-10-374-780A-1897	Sequence 1897, Ap	943	26	81.2	251	16	US-10-425-115-279884	Sequence 279884,
871	26	81.2	223	16	US-10-424-599-268114	Sequence 268114,	944	26	81.2	252	16	US-10-425-115-358451	Sequence 358451,
872	26	81.2	224	16	US-10-437-963-118952	Sequence 118952,	945	26	81.2	252	15	US-10-660-226-18	Sequence 18, Appl
873	26	81.2	225	16	US-09-759-1308-132	Sequence 132, App	946	26	81.2	252	16	US-10-425-115-223026	Sequence 223026,
874	26	81.2	225	16	US-10-741-790-132	Sequence 132, App	947	26	81.2	252	16	US-10-425-115-252358	Sequence 252358,
875	26	81.2	226	15	US-10-282-122A-66135	Sequence 66135, A	948	26	81.2	253	14	US-10-304-946-12	Sequence 12, Appl
876	26	81.2	227	15	US-10-425-114-59804	Sequence 59804, A	949	26	81.2	253	15	US-10-424-599-174188	Sequence 174188,
877	26	81.2	228	15	US-10-425-114-59804	Sequence 68340, A	950	26	81.2	253	16	US-10-800-197-9	Sequence 9, Appli
878	26	81.2	230	10	US-09-984-276-250	Sequence 250, App	951	26	81.2	254	16	US-10-425-115-337669	Sequence 337669,
879	26	81.2	230	10	US-09-984-271-250	Sequence 16, Appl	952	26	81.2	254	17	US-10-890-629-12	Sequence 12, Appl
880	26	81.2	232	14	US-10-112-221A-16	Sequence 16, Appl	953	26	81.2	256	15	US-10-424-599-157900	Sequence 157900,
881	26	81.2	232	15	US-10-424-599-248662	Sequence 248662,	954	26	81.2	256	16	US-10-437-963-119383	Sequence 119383,
882	26	81.2	232	15	US-10-425-114-54241	Sequence 54241, A	955	26	81.2	257	15	US-10-424-599-197883	Sequence 197883,
883	26	81.2	232	16	US-10-437-963-183971	Sequence 183971,	956	26	81.2	258	14	US-10-101-464A-723	Sequence 723, App
884	26	81.2	234	16	US-10-437-963-150323	Sequence 150323,	957	26	81.2	258	16	US-10-432-934-78	Sequence 78, Appl
885	26	81.2	235	10	US-09-533-029-74	Sequence 74, Appl	958	26	81.2	258	16	US-10-437-963-158121	Sequence 158121,
886	26	81.2	235	14	US-10-378-536-104	Sequence 104, App	959	26	81.2	258	17	US-10-864-252-723	Sequence 723, App
887	26	81.2	235	14	US-10-029-586-32032	Sequence 32032, A	960	26	81.2	259	15	US-10-424-599-227542	Sequence 227542,

961 26 81.2 259 15 US-10-424-599-258645 Sequence 258645,
962 26 81.2 259 16 US-10-437-963-116233 Sequence 116233,
963 26 81.2 259 17 US-10-732-923-11974 Sequence 11974, A
964 26 81.2 260 9 US-09-925-300-1458 Sequence 1458, Ap
965 26 81.2 260 15 US-10-424-599-277183 Sequence 277183,
966 26 81.2 260 16 US-10-425-115-338892 Sequence 338892,
967 26 81.2 261 15 US-10-424-599-175955 Sequence 175955,
968 26 81.2 261 16 US-10-437-963-107646 Sequence 107646,
969 26 81.2 261 20 US-11-040-661-24 Sequence 24, Appl
970 26 81.2 262 14 US-10-238-075-1324 Sequence 1324, Ap
971 26 81.2 262 16 US-10-767-701-37335 Sequence 37335, A
972 26 81.2 262 16 US-10-767-701-45810 Sequence 45810, A
973 26 81.2 262 16 US-10-425-115-321421 Sequence 321421,
974 26 81.2 263 15 US-10-425-114-63204 Sequence 63204, A
975 26 81.2 263 16 US-10-739-930-6595 Sequence 6585, Ap
976 26 81.2 265 16 US-10-437-963-186470 Sequence 186470,
977 26 81.2 266 16 US-10-437-963-144000 Sequence 144000,
978 26 81.2 267 15 US-10-282-122A-73006 Sequence 73006, A
979 26 81.2 267 16 US-10-437-963-176506 Sequence 176506,
980 26 81.2 271 14 US-10-267-219-6 Sequence 6, Appl
981 26 81.2 271 15 US-10-156-214A-270 Sequence 270, Appl
982 26 81.2 271 16 US-10-425-115-324070 Sequence 324070,
983 26 81.2 273 14 US-10-195-144-33 Sequence 33, Appl
984 26 81.2 273 15 US-10-345-072-33 Sequence 33, Appl
985 26 81.2 273 16 US-10-437-963-111163 Sequence 111163,
986 26 81.2 273 16 US-10-425-115-317835 Sequence 317835,
987 26 81.2 274 16 US-10-425-115-224046 Sequence 224046,
988 26 81.2 275 16 US-10-437-963-116955 Sequence 116955,
989 26 81.2 275 16 US-10-425-115-220758 Sequence 220758,
990 26 81.2 276 16 US-10-437-963-105809 Sequence 105809,
991 26 81.2 277 9 US-09-965-529-15 Sequence 15, Appl
992 26 81.2 277 10 US-09-969-680A-15 Sequence 15, Appl
993 26 81.2 277 20 US-11-048-692-15 Sequence 15, Appl
994 26 81.2 278 14 US-10-304-946-10 Sequence 10, Appl
995 26 81.2 278 14 US-10-304-946-11 Sequence 11, Appl
996 26 81.2 278 14 US-10-304-946-13 Sequence 13, Appl
997 26 81.2 279 16 US-10-437-963-180700 Sequence 180700,
998 26 81.2 279 16 US-10-767-701-44727 Sequence 44727, A
999 26 81.2 279 16 US-10-767-701-45373 Sequence 45373, A
1000 26 81.2 279 16 US-10-425-115-260567 Sequence 260567,

ALIGNMENTS

RESULT 1
US-10-812-238A-32
; Sequence 32, Application US/10812238A
; Publication No. US20050002904A1
; GENERAL INFORMATION:
; APPLICANT: Wary, Kishore, K.
; APPLICANT: Humtsoe, Joseph O.
; TITLE OF INVENTION: Uses of Vascular Endothelial Growth Factor
; TITLE OF INVENTION: and Type I Collagen Inducible Protein (VCIP)
; FILE REFERENCE: D6563
; CURRENT APPLICATION NUMBER: US/10/812,238A
; CURRENT FILING DATE: 2004-03-29
; PRIOR APPLICATION NUMBER: US 60/458,164
; PRIOR FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 36
; SEQ ID NO 32
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: a peptide containing a RGD sequence
US-10-812-238A-32

Query Match 100.0%; Score 32; DB 17; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.6e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
Db 1 CRGDD 5
RESULT 2
US-10-812-238A-20
; Sequence 20, Application US/10812238A
; Publication No. US20050002904A1
; GENERAL INFORMATION:
; APPLICANT: Wary, Kishore, K.
; APPLICANT: Humtsoe, Joseph O.
; TITLE OF INVENTION: Uses of Vascular Endothelial Growth Factor
; TITLE OF INVENTION: and Type I Collagen Inducible Protein (VCIP)
; FILE REFERENCE: D6563
; CURRENT APPLICATION NUMBER: US/10/812,238A
; CURRENT FILING DATE: 2004-03-29
; PRIOR APPLICATION NUMBER: US 60/458,164
; PRIOR FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 36
; SEQ ID NO 20
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: a peptide containing RGD sequence
US-10-812-238A-20

Query Match 100.0%; Score 32; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
Db 4 CRGDD 8

RESULT 3
US-10-812-238A-23
; Sequence 23, Application US/10812238A
; Publication No. US20050002904A1
; GENERAL INFORMATION:
; APPLICANT: Wary, Kishore, K.
; APPLICANT: Humtsoe, Joseph O.
; TITLE OF INVENTION: Uses of Vascular Endothelial Growth Factor
; TITLE OF INVENTION: and Type I Collagen Inducible Protein (VCIP)
; FILE REFERENCE: D6563
; CURRENT APPLICATION NUMBER: US/10/812,238A
; CURRENT FILING DATE: 2004-03-29
; PRIOR APPLICATION NUMBER: US 60/458,164
; PRIOR FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 36
; SEQ ID NO 23
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: a peptide containing RGD sequence
US-10-812-238A-23

Query Match 100.0%; Score 32; DB 17; Length 13;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
Db 4 CRGDD 8

RESULT 4
US-10-812-238A-2

; Sequence 2, Application US/10812238A
; Publication No. US20050002904A1

; GENERAL INFORMATION:
; APPLICANT: Wary, Kishore, K.
; APPLICANT: Huntsoe, Joseph O.
; TITLE OF INVENTION: Uses of Vascular Endothelial Growth Factor
; TITLE OF INVENTION: and Type I Collagen Inducible Protein (VCIP)
; FILE REFERENCE: D6563
; CURRENT APPLICATION NUMBER: US/10/812,238A
; CURRENT FILING DATE: 2004-03-29
; PRIOR APPLICATION NUMBER: US 60/458,164
; PRIOR FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 36
; SEQ ID NO 2
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: peptide used to raise anti-VCIP-RGD antibody

US-10-812-238A-2

Query Match 100.0%; Score 32; DB 17; Length 20;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 9 CRGDD 13

RESULT 5

US-10-437-963-157113
; Sequence 157113, Application US/10437963

; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 157113
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(43)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_56716C.1.pep

US-10-437-963-157113

Query Match 100.0%; Score 32; DB 16; Length 43;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 30 CRGDD 34

RESULT 6

US-10-425-115-262985

; Sequence 262985, Application US/10425115
; Publication No. US20040214272A1

; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 262985
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_171454C.1.pep

US-10-425-115-262985

Query Match 100.0%; Score 32; DB 16; Length 74;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 44 CRGDD 48

RESULT 7

US-10-437-963-170560

; Sequence 170560, Application US/10437963
; Publication No. US20040123343A1

; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 170560
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_68875C.1.pep

US-10-437-963-170560

Query Match 100.0%; Score 32; DB 16; Length 77;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 36 CRGDD 40

RESULT 8

US-10-437-963-113416

; Sequence 113416, Application US/10437963
; Publication No. US20040123343A1

; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 113416
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_17205C.1.pep
US-10-437-963-113416

Query Match 100.0%; Score 32; DB 16; Length 81;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 35 CRGDD 39

RESULT 9
US-10-389-566-938
; Sequence 938, Application US/10389566
; Publication No. US20040025202A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; APPLICANT: Laurie, Cathy C
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
; FILE REFERENCE: 38-77(52900)D
; CURRENT APPLICATION NUMBER: US/10/389,566
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: US 60/365,301
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/391,786
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/392,018
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 2459
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 938
; LENGTH: 152
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; NAME/KEY: misc_feature
; LOCATION: (88)..(88)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (152)..(152)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-10-389-566-938

Query Match 100.0%; Score 32; DB 15; Length 152;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 49 CRGDD 53

RESULT 10
US-10-425-115-261512
; Sequence 261512, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 261512
; LENGTH: 155
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_170112C.1.pep
US-10-425-115-261512

Query Match 100.0%; Score 32; DB 16; Length 155;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 74 CRGDD 78

RESULT 11
US-10-437-963-160624
; Sequence 160624, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 160624
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_59888C.1.pep
US-10-437-963-160624

Query Match 100.0%; Score 32; DB 16; Length 156;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 35 CRGDD 39

RESULT 12
US-10-424-599-245759
; Sequence 245759, Application US/10424599

; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 245759
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(192)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_63952C.1.pep
US-10-424-599-245759

Query Match 100.0%; Score 32; DB 15; Length 192;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
|
|
|
|
Db 37 CRGDD 41

RESULT 13
US-10-424-599-254290
; Sequence 254290, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 254290
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(203)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_71648C.1.pep
US-10-424-599-254290

Query Match 100.0%; Score 32; DB 15; Length 203;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
|
|
|
|
Db 177 CRGDD 181

RESULT 14
US-10-437-963-199332
; Sequence 199332, Application US/10437963
; Publication No. US20040123343A1

; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 199332
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_94907C.1.pep
US-10-437-963-199332

Query Match 100.0%; Score 32; DB 16; Length 207;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
|
|
|
|
Db 126 CRGDD 130

RESULT 15
US-10-437-963-128934
; Sequence 128934, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 128934
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_3123C.1.pep
US-10-437-963-128934

Query Match 100.0%; Score 32; DB 16; Length 248;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
|
|
|
|
Db 21 CRGDD 25

RESULT 16
US-10-437-963-204891
; Sequence 204891, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:

```
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 204891
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(253)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATUERE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_99933C.1.pap
US-10-437-963-204891

Query Match      100.0%; Score 32; DB 16; Length 253;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CRGDD 5
Db      86 CRGDD 90

RESULT 17
US-10-732-923-5779
; Sequence 5779, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 5779
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(267)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-732-923-5779

Query Match      100.0%; Score 32; DB 17; Length 267;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CRGDD 5
Db      220 CRGDD 224

RESULT 18
US-10-437-963-103596
; Sequence 103596, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
```

```
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 103596
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_101009C.1.pap
US-10-437-963-103596

Query Match      100.0%; Score 32; DB 16; Length 293;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CRGDD 5
Db      43 CRGDD 47

RESULT 19
US-10-369-493-15738
; Sequence 15738, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 15738
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Xanthomonas campestris
US-10-369-493-15738

Query Match      100.0%; Score 32; DB 15; Length 301;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CRGDD 5
Db      200 CRGDD 204

RESULT 20
US-10-369-493-16124
; Sequence 16124, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
```

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493

; CURRENT FILING DATE: 2003-02-28

; PRIOR APPLICATION NUMBER: US 60/360,039

; PRIOR FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 16124

; LENGTH: 301

; TYPE: PRT

; ORGANISM: Xanthomonas campestris

US-10-369-493-16124

Query Match 100.0%; Score 32; DB 15; Length 301;

Best Local Similarity 100.0%; Pred. No. 5.1e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5

Db 200 CRGDD 204

RESULT 21

US-10-369-493-15371

; Sequence 15371, Application US/10369493

; Publication No. US20030233675A1

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Goldman, Barry S.

; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493

; CURRENT FILING DATE: 2003-02-28

; PRIOR APPLICATION NUMBER: US 60/360,039

; PRIOR FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 15371

; LENGTH: 302

; TYPE: PRT

; ORGANISM: Xanthomonas campestris

US-10-369-493-15371

Query Match 100.0%; Score 32; DB 15; Length 302;

Best Local Similarity 100.0%; Pred. No. 5.2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5

Db 203 CRGDD 207

RESULT 22

US-10-655-601-2

; Sequence 2, Application US/10655601

; Publication No. US20040137522A1

; GENERAL INFORMATION:

; APPLICANT: Feany, Mel B.

; APPLICANT: Shulman, Joshua M.

; TITLE OF INVENTION: Genes and Proteins Altering Tau-Related Neuropathy

; FILE REFERENCE: 7570/73251

; CURRENT APPLICATION NUMBER: US/10/655,601

; CURRENT FILING DATE: 2003-09-05

; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 2

; LENGTH: 311

; TYPE: PRT

; ORGANISM: Homo sapiens

Query Match 100.0%; Score 32; DB 16; Length 393;

Best Local Similarity 100.0%; Pred. No. 6.5e+02;

US-10-655-601-2

Query Match 100.0%; Score 32; DB 16; Length 311;

Best Local Similarity 100.0%; Pred. No. 5.3e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5

Db 181 CRGDD 185

RESULT 23

US-10-812-238A-13

; Sequence 13, Application US/10812238A

; Publication No. US20050002904A1

; GENERAL INFORMATION:

; APPLICANT: Wary, Kishore, K.

; APPLICANT: Humtsoe, Joseph O.

; TITLE OF INVENTION: Uses of Vascular Endothelial Growth Factor

; TITLE OF INVENTION: and Type I Collagen Inducible Protein (VCIP)

; FILE REFERENCE: D6563

; CURRENT APPLICATION NUMBER: US/10/812,238A

; CURRENT FILING DATE: 2004-03-29

; PRIOR APPLICATION NUMBER: US 60/458,164

; PRIOR FILING DATE: 2003-03-27

; NUMBER OF SEQ ID NOS: 36

; SEQ ID NO 13

; LENGTH: 311

; TYPE: PRT

; ORGANISM: Unknown

; FEATURE:

; NAME/KEY: CHAIN

; OTHER INFORMATION: human VCIP

US-10-812-238A-13

Query Match 100.0%; Score 32; DB 17; Length 311;

Best Local Similarity 100.0%; Pred. No. 5.3e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5

Db 181 CRGDD 185

RESULT 24

US-10-437-963-132805

; Sequence 132805, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 132805

; LENGTH: 393

; TYPE: PRT

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_34738C.1.pep

US-10-437-963-132805

```
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 175 CRGDD 179

RESULT 25
US-10-282-122A-43293
; Sequence 43293, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43293
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-282-122A-43293

Query Match 100.0%; Score 32; DB 15; Length 400;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 41 CRGDD 45

RESULT 26
US-10-282-122A-59506
; Sequence 59506, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
```

```
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 59506
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-10-282-122A-59506

Query Match 100.0%; Score 32; DB 15; Length 400;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 41 CRGDD 45

RESULT 27
US-10-282-122A-73186
; Sequence 73186, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
```


; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73186
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Salmonella paratyphi A

US-10-282-122A-73186

Query Match 100.0%; Score 32; DB 15; Length 400;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 41 CRGDD 45

RESULT 28

US-10-282-122A-75975
; Sequence 75975, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282.122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75975
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Salmonella typhi

US-10-282-122A-75975

Query Match 100.0%; Score 32; DB 15; Length 400;

Best Local Similarity 100.0%; Pred. No. 6.6e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 41 CRGDD 45

RESULT 29

US-10-425-115-214055
; Sequence 214055, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants

; FILE REFERENCE: 38-21(53222)B

; CURRENT APPLICATION NUMBER: US/10/425.115

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 369326

; SEQ ID NO 214055

; LENGTH: 438

; TYPE: PRT

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: MRT4577_126814C.1.pep

US-10-425-115-214055

Query Match 100.0%; Score 32; DB 16; Length 438;

Best Local Similarity 100.0%; Pred. No. 7.1e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 281 CRGDD 285

RESULT 30

US-10-437-963-103600

; Sequence 103600, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 103600
; LENGTH: 469
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_101012C.1.pep
US-10-437-963-103600

Query Match 100.0%; Score 32; DB 16; Length 469;
Best Local Similarity 100.0%; Pred. No. 7.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 201 CRGDD 205

RESULT 31
US-10-437-963-111036
; Sequence 111036, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 111036
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_15051C.1.pep
US-10-437-963-111036

Query Match 100.0%; Score 32; DB 16; Length 495;
Best Local Similarity 100.0%; Pred. No. 8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 232 CRGDD 236

RESULT 32
US-10-369-493-2069
; Sequence 2069, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 2069
; LENGTH: 1116
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
US-10-369-493-2069

Query Match 100.0%; Score 32; DB 15; Length 1116;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 677 CRGDD 681

RESULT 33
US-10-424-599-257727
; Sequence 257727, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 257727
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_74751C.1.pep
US-10-424-599-257727

Query Match 90.6%; Score 29; DB 15; Length 51;
Best Local Similarity 80.0%; Pred. No. 3.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 13 CRGDD 17

RESULT 34
US-10-424-599-228943
; Sequence 228943, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 228943
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_48762C.1.pep
US-10-424-599-228943

Query Match 90.6%; Score 29; DB 15; Length 62;
Best Local Similarity 80.0%; Pred. No. 4.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
|:|
Db 7 CKGDD 11

RESULT 35

US-10-437-963-124025
; Sequence 124025, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 124025
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_26803C.1.pep
US-10-437-963-124025

Query Match 90.6%; Score 29; DB 16; Length 112;
Best Local Similarity 80.0%; Pred. No. 7.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
|:|
Db 45 CKGDD 49

RESULT 36

US-10-425-115-353920
; Sequence 353920, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 353920
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_85948C.1.pep
US-10-425-115-353920

Query Match 90.6%; Score 29; DB 16; Length 144;
Best Local Similarity 80.0%; Pred. No. 9.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
|:|
Db 53 CKGDD 57

RESULT 37

US-10-425-115-285895
; Sequence 285895, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 285895
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(343)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_23837C.1.pep
US-10-425-115-285895

Query Match 90.6%; Score 29; DB 16; Length 343;
Best Local Similarity 80.0%; Pred. No. 2e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
|:|
Db 218 CKGDD 222

RESULT 38

US-10-424-599-228246
; Sequence 228246, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 228246
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_48134C.1.pep
US-10-424-599-228246

Query Match 90.6%; Score 29; DB 15; Length 360;
Best Local Similarity 80.0%; Pred. No. 2.1e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
|:|
Db 7 CKGDD 11

RESULT 39

US-10-424-599-153174
; Sequence 153174, Application US/10424599
; Publication No. US20040031072A1

```
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 153174
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_109340C.1.pep
US-10-424-599-153174

Query Match          90.6%; Score 29; DB 15; Length 370;
Best Local Similarity 80.0%; Pred. No. 2.2e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CRGDD 5
Db      312 CKGDD 316

RESULT 40
US-10-424-599-151975
; Sequence 151975, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 151975
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(381)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_108257C.1.pep
US-10-424-599-151975

Query Match          90.6%; Score 29; DB 15; Length 381;
Best Local Similarity 80.0%; Pred. No. 2.2e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CRGDD 5
Db      140 CKGDD 144

RESULT 41
US-10-425-115-330473
; Sequence 330473, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
```

```
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 330473
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_64492C.1.pep
US-10-425-115-330473

Query Match          90.6%; Score 29; DB 16; Length 393;
Best Local Similarity 80.0%; Pred. No. 2.3e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CRGDD 5
Db      323 CKGDD 327

RESULT 42
US-10-425-115-194403
; Sequence 194403, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 194403
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_108879C.1.pep
US-10-425-115-194403

Query Match          90.6%; Score 29; DB 16; Length 399;
Best Local Similarity 80.0%; Pred. No. 2.3e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CRGDD 5
Db      370 CKGDD 374

RESULT 43
US-10-183-687-481
; Sequence 481, Application US/10183687
; Publication No. US20030204870A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Allen, William B.
; APPLICANT: Cahoon, Rebecca
; APPLICANT: Epelbaum, Sabine
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Jones, Todd
; APPLICANT: Kinney, Tony
; APPLICANT: Klein, Ted
; APPLICANT: Li, Changjiang
; APPLICANT: Oliveira, Igor Cunha
; APPLICANT: Sakai, Hajime
; APPLICANT: Shen, Bo
```

; APPLICANT: Tarczynski, Mitchell C.
; TITLE OF INVENTION: Alteration Of Oil Traits In Plants
; FILE REFERENCE: BBI458 US NA
; CURRENT APPLICATION NUMBER: US/10/183,687
; PRIOR FILING DATE: 2002-06-27
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 481
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Glycine max
US-10-183-687-481

Query Match 90.6%; Score 29; DB 15; Length 412;
Best Local Similarity 80.0%; Pred. No. 2.4e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 325 CKGDD 329

RESULT 44

US-10-739-930-6640
; Sequence 6640, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: KOVALIC, DAVID K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 6640
; LENGTH: 427
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: Clone ID: ARATH-23APR03-C6231_1.p
US-10-739-930-6640

Query Match 90.6%; Score 29; DB 16; Length 427;
Best Local Similarity 80.0%; Pred. No. 2.5e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 184 CKGDD 188

RESULT 45

US-10-437-963-123766
; Sequence 123766, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 123766

; LENGTH: 483
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_26569C.1.p
US-10-437-963-123766

Query Match 90.6%; Score 29; DB 16; Length 483;
Best Local Similarity 80.0%; Pred. No. 2.8e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 289 CKGDD 293

RESULT 46

US-10-433-757-9
; Sequence 9, Application US/10433757
; Publication No. US20050101529A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; YUE, Henry;
; APPLICANT: AZIMZAI, Yalda; KALLICK, Deborah A.;
; APPLICANT: BAUGHN, Mariah R.; GRIFFIN, Jennifer A.;
; APPLICANT: SHARNAKAR, Anita; LAL, Preeti G.;
; APPLICANT: CHAWLA, Narinder K.; HAFALIA, April J.A.;
; APPLICANT: GANDHI, Ameena R.; AU-YOUNG, Janice K.;
; APPLICANT: ELLIOTT, Vicki S.; RAMKUMAR, Javalaxmi;
; APPLICANT: THANGAVELU, Kavitha; LU, Yan;
; APPLICANT: WARREN, Bridget A.; LU, Dzung Aina M.;
; APPLICANT: LEE, Ernestine A.; TRIBOULEY, Catherine M.;
; APPLICANT: ARVIZO, Chandra S.; DELEGANE, Angele M.;
; APPLICANT: YAO, Monique G.; KHAN, Farrah A.;
; APPLICANT: SANJANWALA, Madhusudan M.
; TITLE OF INVENTION: PROTEIN MODIFICATION AND MAINTENANCE MOLECULES
; FILE REFERENCE: PI-0310 USN
; CURRENT APPLICATION NUMBER: US/10/433,757
; CURRENT FILING DATE: 2003-06-06
; PRIOR APPLICATION NUMBER: PCT/US01/46964
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: US 60/264,623
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: US 60/262,851
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/260,110
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 60/257,803
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: US 60/254,399
; PRIOR FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 508
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7485421CD1
US-10-433-757-9

Query Match 90.6%; Score 29; DB 17; Length 508;
Best Local Similarity 80.0%; Pred. No. 2.9e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 261 CKGDD 265

RESULT 47

US-10-044-205A-14
; Sequence 14, Application US/10044205A

```
; Publication No. US20020123464A1
; GENERAL INFORMATION:
; APPLICANT: KAPPELLER-LIBERMANN, Rosana
; TITLE OF INVENTION: BANDARU, Rajasekhar
; TITLE OF INVENTION: 69087, 15821, and 15418, Methods and Compositions of Human Protei
; TITLE OF INVENTION: Uses Thereof
; FILE REFERENCE: 10147-5201
; CURRENT APPLICATION NUMBER: US/10/044,205A
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/242,428
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: US 60/241,884
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/241,877
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Cyprinus carpio
US-10-044-205A-14

Query Match          90.6%; Score 29; DB 13; Length 549;
Best Local Similarity 80.0%; Pred. No. 3.1e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CRGDD 5
Db      438 CRGDD 442

RESULT 48
US-09-872-153-18
; Sequence 18, Application US/09872153
; Patent No. US20020082207A1
; GENERAL INFORMATION:
; APPLICANT: Hirst, Shannon K.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER
; FILE REFERENCE: 210121.531
; CURRENT APPLICATION NUMBER: US/09/872,153
; CURRENT FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-872-153-18

Query Match          90.6%; Score 29; DB 9; Length 593;
Best Local Similarity 80.0%; Pred. No. 3.3e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CRGDD 5
Db      346 CRGDD 350

RESULT 49
US-10-723-860-3727
; Sequence 3727, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Nataasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
```

```
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3727
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-723-860-3727

Query Match          90.6%; Score 29; DB 16; Length 593;
Best Local Similarity 80.0%; Pred. No. 3.3e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CRGDD 5
Db      346 CRGDD 350

RESULT 50
US-10-437-963-121532
; Sequence 121532, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 121532
; LENGTH: 813
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURES:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_24548C.1.pep
US-10-437-963-121532

Query Match          90.6%; Score 29; DB 16; Length 813;
Best Local Similarity 80.0%; Pred. No. 4.3e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CRGDD 5
Db      17 CRGDD 21

RESULT 51
US-09-888-615-66
; Sequence 66, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: FLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEFEEL, SEAN
; APPLICANT: CHARYDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
```

; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 66
; LENGTH: 953
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-615-66

Query Match 90.6%; Score 29; DB 9; Length 953;
Best Local Similarity 80.0%; Pred. No. 5e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 695 CKGDD 699

RESULT 52

US-10-311-035-21
; Sequence 21, Application US/10311035
; Publication No. US20040023243A1

; GENERAL INFORMATION:

; APPLICANT: INCYTE GENOMICS, INC.

; APPLICANT: YUE, Henry

; APPLICANT: ELLIOTT, Vicki

; APPLICANT: GANDHI, Ameena R.

; APPLICANT: LAL, Preeti

; APPLICANT: AU-YOUNG, Janice

; APPLICANT: TRIBOULEY, Catherine M.

; APPLICANT: DELEGEANE, Angelo M.

; APPLICANT: BAUGHN, Mariah R.

; APPLICANT: NGUYEN, Daniel B.

; APPLICANT: LEE, Ernestine A.

; APPLICANT: HAFALIA, April

; APPLICANT: KHAN, Farrah A.

; APPLICANT: CHAWLA, Narinder K.

; APPLICANT: YAO, Monique G.

; APPLICANT: LU, Dyung Aina M.

; APPLICANT: ARVIZU, Chandra S.

; APPLICANT: TANG, Y. Tom

; APPLICANT: WALSH, Roderick T.

; APPLICANT: AZIMZAI, Yalda

; APPLICANT: LU, Yan

; APPLICANT: RAMKUMAR, Jayalaximi

; APPLICANT: XU, Yuming

; APPLICANT: REDDY, Roopa

; APPLICANT: DAS, Depopriya

; APPLICANT: KEARNEY, Liam

; APPLICANT: KALLICK, Deborah A.

; TITLE OF INVENTION: Proteases

; FILE REFERENCE: PI-0123 PCT

; CURRENT APPLICATION NUMBER: US/10/311,035

; CURRENT FILING DATE: 2002-12-10

; PRIOR APPLICATION NUMBER: 60/212,336; 60/213,955; 60/215,396; 60/216,821; 60/218,946

; PRIOR FILING DATE: 2000-06-16; 2000-06-22; 2000-06-29; 2000-07-07; 2000-07-14

; NUMBER OF SEQ ID NOS: 42

; SOFTWARE: PERL Program

; SEQ ID NO 21

; LENGTH: 953

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No. US20040023243A1 6817347CD1

US-10-311-035-21

Query Match 90.6%; Score 29; DB 15; Length 953;
Best Local Similarity 80.0%; Pred. No. 5e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5

Db 695 CKGDD 699

RESULT 53

US-10-437-963-123768

; Sequence 123768, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 123768

; LENGTH: 988

; TYPE: PRT

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_26570C.1.pep

US-10-437-963-123768

Query Match 90.6%; Score 29; DB 16; Length 988;

Best Local Similarity 80.0%; Pred. No. 5.2e+03;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5

Db 601 CKGDD 605

RESULT 54

US-08-865-579-74

; Sequence 74, Application US/08865579

; Publication No. US20010006779A1

; GENERAL INFORMATION:

; APPLICANT: Alnemri, Emad S.

; APPLICANT: Fernandes-Alnemri, Tereesa

; APPLICANT: Litwack, Gerald

; TITLE OF INVENTION: Apoptotic Protease Mch6, Nucleic Acids

; TITLE OF INVENTION: Encoding Same and Methods of Use

; NUMBER OF SEQUENCES: 87

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Campbell & Flores LLP

; STREET: 4370 La Jolla Village Drive, Suite 700

; CITY: San Diego

; STATE: California

; COUNTRY: United States

; ZIP: 92122

; COMPUTER READABLE FORM:

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/865,579

; FILING DATE: 29-MAY-1997

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Campbell, Cathryn A.

; REGISTRATION NUMBER: 31,815

; REFERENCE/DOCKET NUMBER: P-ID 2180

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (619) 535-9001

Qy 1 CRGDD 5

TELEFAX: (619) 535-9849
; INFORMATION FOR SEQ ID NO: 74:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-865-579-74

Query Match 87.5%; Score 28; DB 8; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.6e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
Db 5 CRGDE 9

RESULT 55
US-09-746-731-74
; Sequence 74, Application US/09746731
; Publication No. US20010016345A1
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
; Fernandes-Alnemri, Teresa
; Litwack, Gerald
; TITLE OF INVENTION: Apoptotic Protease Mch6, Nucleic Acids
; NUCLEIC ACIDS ENCODING AND METHODS OF USE
; NUMBER OF SEQUENCES: 87
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/746.731
; FILING DATE: 22-Dec-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/865,579
; FILING DATE: 29-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-ID 2180
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9849
; TELEFAX: (619) 535-9849
; INFORMATION FOR SEQ ID NO: 74:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 74:
US-09-746-731-74

Query Match 87.5%; Score 28; DB 9; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.6e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
Db 5 CRGDE 9

RESULT 56
US-09-952-768-59
; Sequence 59, Application US/09952768
; Patent No. US20020035242A1
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
; Fernandes-Alnemri, Teresa
; Litwack, Gerald
; Armstrong, Robert
; Tomaselli, Kevin
; TITLE OF INVENTION: MCH4 AND MCH5, APOPTOTIC PROTEASE,
; NUCLEIC ACIDS ENCODING AND METHODS OF USE
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed Intellectual Property Law Group
; STREET: Suite 6300, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/952,768
; FILING DATE: 10-Sep-2001
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Christiansen, William T.
; REGISTRATION NUMBER: 44,614
; REFERENCE/DOCKET NUMBER: 480140.424C4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..9
; OTHER INFORMATION: /note= "ICH-1"
; SEQUENCE DESCRIPTION: SEQ ID NO: 59:
US-09-952-768-59

Query Match 87.5%; Score 28; DB 9; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.6e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
Db 5 CRGDE 9

RESULT 57
US-09-954-697-112
; Sequence 112, Application US/09954697
; Patent No. US20020106631A1
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
; TITLE OF INVENTION: RECOMBINANT, ACTIVE CASPASES AND USSES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 480140.431D2
; CURRENT APPLICATION NUMBER: US/09/954,697
; CURRENT FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 112


```
;
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-954-697-112

Query Match      87.5%; Score 28; DB 9; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.6e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CRGDD 5
Db 5 CRGDE 9

RESULT 58
US-10-059-749-74
; Sequence 74, Application US/10059749
; Publication No. US20020183504A1
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
;               Fernandes-Alnemri, Teresa
;               Litwack, Gerald
; TITLE OF INVENTION: Apoptotic Protease Mch6, Nucleic Acids
;               Encoding Same and Methods of Use
; NUMBER OF SEQUENCES: 87
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/059,749
; FILING DATE: 29-Jan-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/865,579
; FILING DATE: 29-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-ID 2180
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-9849
; INFORMATION FOR SEQ ID NO: 74:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 74:
US-10-059-749-74

Query Match      87.5%; Score 28; DB 13; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.6e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CRGDD 5
Db 5 CRGDE 9

RESULT 59
US-10-668-955-59
; Sequence 59, Application US/10668955
; Publication No. US20040054148A1

; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
;               Fernandes-Alnemri, Teresa
;               Litwack, Gerald
;               Armstrong, Robert
;               Tomaselli, Kevin
; TITLE OF INVENTION: MCH4 AND MCH5, APOPTOTIC PROTEASE,
;               NUCLEIC ACIDS ENCODING AND METHODS OF USE
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed Intellectual Property Law Group
; STREET: Suite 6300, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/668,955
; FILING DATE: 22-Sep-2003
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Laherty, Carol D.
; REGISTRATION NUMBER: 51,909
; REFERENCE/DOCKET NUMBER: 480140.424D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..9
; OTHER INFORMATION: /note= "ICH-1"
; SEQUENCE DESCRIPTION: SEQ ID NO: 59:
US-10-668-955-59

Query Match      87.5%; Score 28; DB 15; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.6e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CRGDD 5
Db 5 CRGDE 9

RESULT 60
US-10-812-238A-22
; Sequence 22, Application US/10812238A
; Publication No. US20050002904A1
; GENERAL INFORMATION:
; APPLICANT: Wary, Kishore, K.
; APPLICANT: Humtsoe, Joseph O.
; TITLE OF INVENTION: Uses of Vascular Endothelial Growth Factor
; TITLE OF INVENTION: and Type I Collagen Inducible Protein (VCIP)
; FILE REFERENCE: D6563
; CURRENT APPLICATION NUMBER: US/10/812,238A
; CURRENT FILING DATE: 2004-03-29
; PRIOR APPLICATION NUMBER: US 60/458,164
; PRIOR FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 36
; SEQ ID NO 22
; LENGTH: 10
; TYPE: PRT
```

; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: a peptide containing a mutated RGD sequence
US-10-812-238A-22

Query Match 87.5%; Score 28; DB 17; Length 10;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
|||:
Db 4 CRGED 8

RESULT 61
US-10-133-128-62
; Sequence 62, Application US/10133128
; Publication No. US20030082630A1
; GENERAL INFORMATION:
; APPLICANT: KOLKMAN, JOOST A.
; TITLE OF INVENTION: COMBINATORIAL LIBRARIES OF MONOMER DOMAINS
; FILE REFERENCE: 0319.410US
; CURRENT APPLICATION NUMBER: US/10/133,128
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: 60/374,107
; PRIOR FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: 60/333,359
; PRIOR FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: 60/337,209
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/286,823
; PRIOR FILING DATE: 2001-04-26
; NUMBER OF SEQ ID NOS: 244
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-133-128-62

Query Match 87.5%; Score 28; DB 14; Length 35;
Best Local Similarity 80.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
|||:
Db 1 CRGDE 5

RESULT 62
US-10-289-660-62
; Sequence 62, Application US/10289660
; Publication No. US20030157561A1
; GENERAL INFORMATION:
; APPLICANT: KOLKMAN, JOOST A.
; APPLICANT: STEMMER, WILLEM P.C.
; TITLE OF INVENTION: COMBINATORIAL LIBRARIES OF MONOMER DOMAINS
; FILE REFERENCE: 0319.510US
; CURRENT APPLICATION NUMBER: US/10/289,660
; CURRENT FILING DATE: 2003-11-06
; PRIOR APPLICATION NUMBER: 10/133,128
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: 60/374,107
; PRIOR FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: 60/333,359
; PRIOR FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: 60/337,209
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/286,823
; PRIOR FILING DATE: 2001-04-26

; NUMBER OF SEQ ID NOS: 244
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-289-660-62

Query Match 87.5%; Score 28; DB 14; Length 35;
Best Local Similarity 80.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
|||:
Db 1 CRGDE 5

RESULT 63
US-10-693-057-62
; Sequence 62, Application US/10693057
; Publication No. US20040175756A1
; GENERAL INFORMATION:
; APPLICANT: Kolkman, Joost A.
; APPLICANT: Stemmer, Willem P.C.
; APPLICANT: Avidia Research Institute
; TITLE OF INVENTION: Methods for Using Combinatorial Libraries
; TITLE OF INVENTION: of Monomer Domains
; FILE REFERENCE: 022013-000170US
; CURRENT APPLICATION NUMBER: US/10/693,057
; CURRENT FILING DATE: 2003-10-24
; PRIOR APPLICATION NUMBER: US 60/286,823
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: US 60/337,209
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: US 60/333,359
; PRIOR FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: US 60/374,107
; PRIOR FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: US 10/133,128
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: US 10/289,660
; PRIOR FILING DATE: 2002-11-06
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 62
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: human ApoER2 A domain
US-10-693-057-62

Query Match 87.5%; Score 28; DB 16; Length 35;
Best Local Similarity 80.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
|||:
Db 1 CRGDE 5

RESULT 64
US-10-693-056-62
; Sequence 62, Application US/10693056
; Publication No. US20050048512A1
; GENERAL INFORMATION:
; APPLICANT: Kolkman, Joost A.
; APPLICANT: Stemmer, Willem P.C.
; APPLICANT: Avidia Research Institute
; TITLE OF INVENTION: Combinatorial Libraries of Monomer Domains
; FILE REFERENCE: 022013-000160US
; CURRENT APPLICATION NUMBER: US/10/693,056
; CURRENT FILING DATE: 2003-10-24

; PRIOR APPLICATION NUMBER: US 60/286,823
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: US 60/337,209
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: US 60/333,359
; PRIOR FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: US 60/374,107
; PRIOR FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: US 10/133,128
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: US 10/289,660
; PRIOR FILING DATE: 2002-11-06
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 62
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: human ApoER2 A domain
US-10-693-056-62

Query Match 87.5%; Score 28; DB 17; Length 35;
Best Local Similarity 80.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 1 CRGDE 5

RESULT 65
US-10-840-723-62
; Sequence 62, Application US/10840723
; Publication No. US2005053973A1
; GENERAL INFORMATION:
; APPLICANT: Kolkman, Joost A.
; APPLICANT: Stemmer, Willem P. C.
; APPLICANT: Freskgard, Per-Ola
; APPLICANT: Avidia Research Institute
; TITLE OF INVENTION: Novel Proteins With Targeted Binding
; FILE REFERENCE: 022013-000180US
; CURRENT APPLICATION NUMBER: US/10/840,723
; PRIOR FILING DATE: 2004-05-05
; PRIOR APPLICATION NUMBER: US 60/286,823
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: US 60/337,209
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: US 60/333,359
; PRIOR FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: US 60/374,107
; PRIOR FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: US 10/133,128
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: US 10/289,660
; PRIOR FILING DATE: 2002-11-06
; PRIOR APPLICATION NUMBER: US 10/693,056
; PRIOR FILING DATE: 2003-10-24
; PRIOR APPLICATION NUMBER: US 10/693,057
; PRIOR FILING DATE: 2003-10-24
; NUMBER OF SEQ ID NOS: 665
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 62
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: human ApoER2 A domain
US-10-840-723-62

Query Match 87.5%; Score 28; DB 17; Length 35;
Best Local Similarity 80.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 1 CRGDE 5

RESULT 66
US-10-871-602-62
; Sequence 62, Application US/10871602
; Publication No. US2005089932A1
; GENERAL INFORMATION:
; APPLICANT: Kolkman, Joost A.
; APPLICANT: Stemmer, Willem P. C.
; APPLICANT: Freskgard, Per-Ola
; APPLICANT: Avidia Research Institute
; TITLE OF INVENTION: Novel Proteins With Targeted Binding
; FILE REFERENCE: 022013-000190US
; CURRENT APPLICATION NUMBER: US/10/871,602
; CURRENT FILING DATE: 2004-06-17
; PRIOR APPLICATION NUMBER: US 60/286,823
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: US 60/337,209
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: US 60/333,359
; PRIOR FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: US 60/374,107
; PRIOR FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: US 10/133,128
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: US 10/289,660
; PRIOR FILING DATE: 2002-11-06
; NUMBER OF SEQ ID NOS: 665
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 62
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: human ApoER2 A domain
US-10-871-602-62

Query Match 87.5%; Score 28; DB 17; Length 35;
Best Local Similarity 80.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 1 CRGDE 5

RESULT 67
US-10-971-679-62
; Sequence 62, Application US/10971679
; Publication No. US20050164301A1
; GENERAL INFORMATION:
; APPLICANT: Kolkman, Joost A.
; APPLICANT: Stemmer, Willem P. C.
; APPLICANT: Avidia Research Institute
; TITLE OF INVENTION: LDL Receptor Class A and EGF Domain
; FILE REFERENCE: 022013-000210US
; CURRENT APPLICATION NUMBER: US/10/971,679
; CURRENT FILING DATE: 2004-10-22
; PRIOR APPLICATION NUMBER: US 60/514,391
; PRIOR FILING DATE: 2003-10-24
; PRIOR APPLICATION NUMBER: US 10/693,056
; PRIOR FILING DATE: 2003-10-24
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 62
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Artificial Sequence

```
;
; FEATURE:
; OTHER INFORMATION: human ApoER2 A domain
US-10-971-679-62

Query Match      87.5%; Score 28; DB 18; Length 35;
Best Local Similarity 80.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 1 CRGDE 5

RESULT 68
US-11-066-697-1545
; Sequence 1545, Application US/11066697
; Publication No. US20050187159A1
; GENERAL INFORMATION:
; APPLICANT: Bridon, Dominique P.
; APPLICANT: Ezrin, Alan M.
; APPLICANT: Milner, Peter G.
; APPLICANT: Holmes, Darren L.
; APPLICANT: Thibaudau, Karen
; TITLE OF INVENTION: PROTECTION OF ENDOGENOUS THERAPEUTIC PEPTIDES FROM
; TITLE OF INVENTION: PEPTIDASE ACTIVITY THROUGH CONJUGATION TO BLOOD
; TITLE OF INVENTION: COMPONENTS
; FILE REFERENCE: 500862002301
; CURRENT APPLICATION NUMBER: US/11/066,697
; CURRENT FILING DATE: 2005-02-25
; PRIOR APPLICATION NUMBER: 09/657,276
; PRIOR FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 60/153,406
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: 60/159,783
; PRIOR FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 1617
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1545
; LENGTH: 39
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-11-066-697-1545

Query Match      87.5%; Score 28; DB 20; Length 39;
Best Local Similarity 80.0%; Pred. No. 4.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 7 CQGDD 11

RESULT 69
US-08-865-579-16
; Sequence 16, Application US/08865579
; Publication No. US20010006779A1
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
; APPLICANT: Fernandes-Alnemri, Teresa
; APPLICANT: Litwack, Gerald
; TITLE OF INVENTION: Apoptotic Protease Mch6, Nucleic Acids
; TITLE OF INVENTION: Encoding Same and Methods of Use
; NUMBER OF SEQUENCES: 87
; CORRESPONDENCE ADDRESS:
; ADDRESSSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
```

```
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/865,579
; FILING DATE: 29-MAY-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-ID 2180
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-9849
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-865-579-16

Query Match      87.5%; Score 28; DB 8; Length 46;
Best Local Similarity 80.0%; Pred. No. 5.4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 17 CRGDE 21

RESULT 70
US-09-746-731-16
; Sequence 16, Application US/09746731
; Publication No. US20010016345A1
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
; APPLICANT: Fernandes-Alnemri, Teresa
; APPLICANT: Litwack, Gerald
; TITLE OF INVENTION: Apoptotic Protease Mch6, Nucleic Acids
; TITLE OF INVENTION: Encoding Same and Methods of Use
; NUMBER OF SEQUENCES: 87
; CORRESPONDENCE ADDRESS:
; ADDRESSSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/746,731
; FILING DATE: 22-Dec-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/865,579
; FILING DATE: 29-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-ID 2180
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-9849
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46 amino acids
```

```
;
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-09-746-731-16
Query Match      87.5%; Score 28; DB 9; Length 46;
Best Local Similarity 80.0%; Pred. No. 5.4e+02; Indels 0; Gaps 0;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CRGDD 5
Db      17 CRGDE 21

RESULT 71
US-10-059-749-16
; Sequence 16, Application US/10059749
; Publication No. US20020183504A1
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
;               Fernandes-Alnemri, Teresa
;               Litwack, Gerald
; TITLE OF INVENTION: Apoptotic Protease Mch6, Nucleic Acids
;               Encoding Same and Methods of Use
; NUMBER OF SEQUENCES: 87
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/059,749
; FILING DATE: 29-Jan-2002
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/865,579
; FILING DATE: 29-MAY-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/POCKET NUMBER: P-ID 2180
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-9849
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-10-059-749-16
Query Match      87.5%; Score 28; DB 13; Length 46;
Best Local Similarity 80.0%; Pred. No. 5.4e+02; Indels 0; Gaps 0;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CRGDD 5
Db      17 CRGDE 21

RESULT 72
US-09-989-903-58
; Sequence 58, Application US/09989903
; Patent No. US20020146804A1
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
; TITLE OF INVENTION: CASPASE-14, AN APOTOTIC PROTEASE, NUCLEIC ACID ENCODING
; FILE REFERENCE: 480140.434D1
; CURRENT APPLICATION NUMBER: US/09/989,903
; CURRENT FILING DATE: 2002-04-11
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 58
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-989-903-58
Query Match      87.5%; Score 28; DB 9; Length 56;
Best Local Similarity 80.0%; Pred. No. 6.4e+02; Indels 0; Gaps 0;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CRGDD 5
Db      31 CRGDE 35

RESULT 73
US-10-068-564-58
; Sequence 58, Application US/10068564
; Publication No. US20030040096A1
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
; TITLE OF INVENTION: CASPASE-14, AN APOTOTIC PROTEASE, NUCLEIC ACID ENCODING
; FILE REFERENCE: 480140.434C2
; CURRENT APPLICATION NUMBER: US/10/068,564
; CURRENT FILING DATE: 2002-02-05
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 58
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-068-564-58
Query Match      87.5%; Score 28; DB 14; Length 56;
Best Local Similarity 80.0%; Pred. No. 6.4e+02; Indels 0; Gaps 0;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CRGDD 5
Db      31 CRGDE 35

RESULT 74
US-10-870-765-58
; Sequence 58, Application US/10870765
; Publication No. US20040236081A1
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
; TITLE OF INVENTION: CASPASE-14, AN APOTOTIC PROTEASE, NUCLEIC ACID ENCODING
; FILE REFERENCE: 480140.434D2
; CURRENT APPLICATION NUMBER: US/10/870,765
; CURRENT FILING DATE: 2004-06-17
; PRIOR APPLICATION NUMBER: 10/068,564
; PRIOR FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 09/989,903
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/187,789
; PRIOR FILING DATE: 1998-11-06
```

; PRIOR APPLICATION NUMBER: 09/139,600
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 58
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-870-765-58

Query Match 87.5%; Score 28; DB 16; Length 56;
Best Local Similarity 80.0%; Pred. No. 6.4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 31 CRGDE 35

RESULT 75
US-10-767-701-54704
; Sequence 54704, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 54704
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 14592790.pcp
US-10-767-701-54704

Query Match 87.5%; Score 28; DB 16; Length 72;
Best Local Similarity 80.0%; Pred. No. 7.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 38 CRGDE 42

Search completed: September 7, 2005, 20:13:46
Job time : 176 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2005, 19:57:40 ; Search time 163 Seconds
(without alignments)
11.864 Million cell updates/sec

Title: US-10-812-238B-41

Perfect score: 32

Sequence: 1 CRGDD 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : A_Geneseq_16Dec04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	32	100.0	7	2	AAR96393
2	32	100.0	7	2	AAR66831
3	32	100.0	7	2	AAR96984
4	32	100.0	9	2	AAR96408
5	32	100.0	9	2	AAR66840
6	32	100.0	9	2	AAR97003
7	32	100.0	92	4	AAR61299
8	32	100.0	92	6	ABM57818
9	32	100.0	117	6	ABR41549
10	32	100.0	120	4	AAU44177
11	32	100.0	120	6	ABM40696
12	32	100.0	152	8	ADU348934
13	32	100.0	154	3	AAB40710
14	32	100.0	156	7	ADC39079
15	32	100.0	301	8	ADS26705
16	32	100.0	301	8	ADS27091
17	32	100.0	302	8	ADS26338
18	32	100.0	311	2	AAR79286
19	32	100.0	311	7	ADBE60703
20	32	100.0	311	8	ADM72103
21	32	100.0	311	8	ADN04415
22	32	100.0	311	8	ADP24274
23	32	100.0	376	7	ABO62272
24	32	100.0	400	6	ABU31582
25	32	100.0	400	6	ABU45262

26	32	100.0	400	6	ABU48051
27	32	100.0	400	6	ABU15369
28	32	100.0	424	7	ABO65513
29	32	100.0	488	7	ABO66647
30	32	100.0	1116	8	ADN19416
31	32	100.0	1837	4	ABB62913
32	29	90.6	7	2	AAR96407
33	29	90.6	7	2	AAR66822
34	29	90.6	41	4	AAR97002
35	29	90.6	105	7	ABO67045
36	29	90.6	274	6	ADB08040
37	29	90.6	299	8	ADJ71969
38	29	90.6	357	6	ADA50457
39	29	90.6	412	6	ABR40851
40	29	90.6	435	8	ADJ16354
41	29	90.6	508	5	ABB98131
42	29	90.6	514	7	ADF92397
43	29	90.6	525	4	ABG10394
44	29	90.6	525	8	ADH54836
45	29	90.6	525	8	ABM81830
46	29	90.6	528	4	AAU23205
47	29	90.6	556	7	ADC10088
48	29	90.6	575	7	ADC10090
49	29	90.6	581	7	ADC10086
50	29	90.6	593	5	AAE17560
51	29	90.6	593	8	ADH54838
52	29	90.6	593	8	ADQ20907
53	29	90.6	639	7	ABO70157
54	29	90.6	936	6	ABU92027
55	29	90.6	953	5	AAU82708
56	29	90.6	953	5	AAU74761
57	29	90.6	9	2	AAR96415
58	28	87.5	9	2	AAR66841
59	28	87.5	9	2	AAU21792
60	28	87.5	9	2	AAU97010
61	28	87.5	9	4	AAE09007
62	28	87.5	9	5	ABU01296
63	28	87.5	21	8	ADM96249
64	28	87.5	35	6	ABU61264
65	28	87.5	35	8	ADP21486
66	28	87.5	39	2	AAR13677
67	28	87.5	39	2	AAR95440
68	28	87.5	39	4	AAB92369
69	28	87.5	41	2	AAR95442
70	28	87.5	41	2	AAW20043
71	28	87.5	42	2	AAR95441
72	28	87.5	42	2	AAR20042
73	28	87.5	45	2	AAR13676
74	28	87.5	46	4	AAE08949
75	28	87.5	46	8	ADE52034
76	28	87.5	55	4	AAU41390
77	28	87.5	55	6	ABM37909
78	28	87.5	57	4	AAU59258
79	28	87.5	57	6	ABM55777
80	28	87.5	59	4	AAU67021
81	28	87.5	65	4	AAO12454
82	28	87.5	65	6	ABM63540
83	28	87.5	66	8	ADG74135
84	28	87.5	69	4	ADM92308
85	28	87.5	74	8	ADG74134
86	28	87.5	86	3	AAG19855
87	28	87.5	86	4	AAB48063
88	28	87.5	86	8	ADL66887
89	28	87.5	93	4	AAU49607
90	28	87.5	93	6	ABM46126
91	28	87.5	105	6	ABM51089
92	28	87.5	106	3	AAG19854
93	28	87.5	107	3	ADQ65061
94	28	87.5	110	3	AAG19853
95	28	87.5	131	4	ABG53535
96	28	87.5	131	4	ABG53535
97	28	87.5	131	4	ABG53535
98	28	87.5	131	4	ABG53535

99	28	87.5	131	5	ABG41664	Abg41664 Human pep	172	28	87.5	4	AAG92510	Aag92510 C glutami
100	28	87.5	137	4	AAU51527	Aau51527 Propionib	173	28	87.5	7	ADM03823	Adm03823 Human pro
101	28	87.5	137	6	ABM48046	Abm48046 Propionib	174	28	87.5	629	ADR31370	Adr31370 Sresstre
102	28	87.5	140	4	ABE11487	Abb11487 Human sec	175	28	87.5	661	ADH71746	Adh71746 Human pro
103	28	87.5	160	4	AAH71915	Aab71915 Human ISO	176	28	87.5	686	ABO69369	Abo69369 Pseudomon
104	28	87.5	160	5	ADH48812	Adh48812 NOV42A pr	177	28	87.5	695	AAU91286	Aau91286 Human NOV
105	28	87.5	160	5	ADH48814	Adh48814 NOV42B pr	178	28	87.5	695	ADH71752	Adh71752 Human pro
106	28	87.5	181	4	ABE67977	Abb67977 Drosophil	179	28	87.5	699	ABU56579	Abu56579 Lung can
107	28	87.5	192	7	ABO66885	Aboc66885 Klebsiell	180	28	87.5	699	ADL06561	Adl06561 Human tum
108	28	87.5	230	4	AAU53393	Abm53393 Propionib	181	28	87.5	700	ADQ26075	Adq26075 Low densi
109	28	87.5	230	6	ABM49912	Abm49912 Propionib	182	28	87.5	722	ABR39968	AbR39968 Human LGR
110	28	87.5	235	8	ADR10154	Adr10154 Human pro	183	28	87.5	724	ABE63834	AbE63834 Drosophi1
111	28	87.5	312	5	ABR99112	AbB99112 Rat phosp	184	28	87.5	729	AAU78665	Aau78665 Human NOV
112	28	87.5	312	7	ADJ71981	Adj71981 Human PWM	185	28	87.5	729	AAU91282	Aau91282 Human NOV
113	28	87.5	323	4	AAU16940	Ade06701 Rat Prote	186	28	87.5	730	AAU91282	Aau91282 Human NOV
114	28	87.5	337	2	AAU36854	Aay36854 Protein i	187	28	87.5	762	AAU91283	Aau91283 Human NOV
115	28	87.5	362	8	ADQ15100	Adq15100 Human can	188	28	87.5	762	AAU78666	Aau78666 Human NOV
116	28	87.5	367	3	ABG39030	Abg39030 Human sec	189	28	87.5	775	ADQ93398	AdQ93398 Human lip
117	28	87.5	379	8	ADJ71981	Adj71981 Human PWM	190	28	87.5	775	ADH71760	Adh71760 Human pro
118	28	87.5	390	7	ADM06034	Adm06034 Human pro	191	28	87.5	800	ABO60674	Aboc60674 Klebsiell
119	28	87.5	392	3	AGC28027	Aag28027 Arabidops	192	28	87.5	804	AAU91287	Aau91287 Human NOV
120	28	87.5	405	6	ABM69291	Abm69291 Photornab	193	28	87.5	804	ADH71754	Adh71754 Human pro
121	28	87.5	409	3	AGC28026	Aag28026 Arabidops	194	28	87.5	809	AAH38157	Aae38157 Fruit fly
122	28	87.5	416	3	AAW03147	Aaw03147 Epstein-B	195	28	87.5	822	AAW99087	Aaw99087 Human ser
123	28	87.5	416	3	AAW98164	Aay98164 Human TPA	196	28	87.5	825	AAU91284	Aau91284 Human NOV
124	28	87.5	416	7	ADF76476	Adf76476 Novel hum	197	28	87.5	825	ADH71748	Adh71748 Human pro
125	28	87.5	416	8	ADL22912	Adl22912 Human MP2	198	28	87.5	834	ADH71748	Adh71748 Human pro
126	28	87.5	416	8	ADL82947	Adl82947 Human PRO	199	28	87.5	834	ADH71762	Adh71762 Human pro
127	28	87.5	416	8	ADO20499	Ado20499 Human PRO	200	28	87.5	842	AAU36357	Aau36357 Pseudomon
128	28	87.5	416	8	ADO19524	Ado19524 Human PRO	201	28	87.5	842	ABU38620	Abu38620 Protein e
129	28	87.5	416	8	ADP56102	Adp56102 Human PRO	202	28	87.5	847	AAU91289	Aau91289 Human NOV
130	28	87.5	416	8	ADS88159	Ades88159 Human pro	203	28	87.5	847	ADH71758	Adh71758 Human pro
131	28	87.5	417	3	AAW87785	Aay87785 Human TNP	204	28	87.5	857	AAU91288	Aau91288 Human NOV
132	28	87.5	417	7	ABW02030	Abw02030 Human TNP	205	28	87.5	857	ADH71756	Adh71756 Human pro
133	28	87.5	417	7	ADG47653	Adg47653 Human TNP	206	28	87.5	875	AAW83361	Aaw83361 Human neu
134	28	87.5	426	7	ABO76570	Aboc76570 Pseudomon	207	28	87.5	875	ADQ21226	AdQ21226 Human sof
135	28	87.5	433	7	ABO61626	Aboc61626 Klebsiell	208	28	87.5	875	ADQ89082	AdQ89082 Human uro
136	28	87.5	435	2	AAE66771	Aar66771 Human int	209	28	87.5	904	ADH71768	Adh71768 Human pro
137	28	87.5	435	2	AAE98462	Aar98462 Human Ice	210	28	87.5	904	ADH71768	Adh71768 Human pro
138	28	87.5	435	2	AAE90703	Aar90703 Interleuk	211	28	87.5	905	AAU91290	Aau91290 Human NOV
139	28	87.5	435	2	AAW26274	Aaw26274 Cell deat	212	28	87.5	905	ADH71742	Adh71742 Human pro
140	28	87.5	435	2	AAW21716	Aay21716 Amino aci	213	28	87.5	905	ADH71766	Adh71766 Human pro
141	28	87.5	435	3	AAE14257	Aab14257 Human Ich	214	28	87.5	963	ABP56838	Abp56838 Human apo
142	28	87.5	435	3	AAE00599	Aae00599 Human cas	215	28	87.5	963	ADH71764	Adh71764 Human pro
143	28	87.5	435	5	ABE78321	Abb78321 Amino aci	216	28	87.5	963	ADI27185	Adi27185 Human lrp
144	28	87.5	435	5	ABJ01217	Abj01217 Human cas	217	28	87.5	963	ADN00737	Adn00737 Human LDL
145	28	87.5	435	7	ADG63002	Ades63002 Human PRO	218	28	87.5	963	ADO19504	Ado19504 Human PRO
146	28	87.5	435	7	ADI63132	Adi63132 Human apo	219	28	87.5	963	ADQ26074	AdQ26074 Low dens1
147	28	87.5	435	8	ADO16842	Ado16842 Human cas	220	28	87.5	1012	AAU91285	Aau91285 Human NOV
148	28	87.5	441	2	AAE66768	Aar66768 Human int	221	28	87.5	1012	ADH71750	Adh71750 Human pro
149	28	87.5	441	3	AAE14253	Aab14253 Human Ich	222	28	87.5	1286	ABG10455	Abg10455 Novel hum
150	28	87.5	441	3	AGC28025	Aag28025 Arabidops	223	28	87.5	1342	ABG01726	Abg01726 Novel hum
151	28	87.5	442	8	ADO26843	Adoc26843 Human rec	224	28	87.5	1806	ADN22969	Adn22969 Bacterial
152	28	87.5	446	5	ABG79606	Abg79606 Wheat rib	225	28	87.5	2257	ABE61782	AbE61782 Drosophi1
153	28	87.5	450	4	ABE63682	Aboc63682 Drosophil	226	28	87.5	3115	ABE59485	AbE59485 Drosophi1
154	28	87.5	451	3	AAE07582	Aab07582 Protein e	227	28	87.5	7	AAE96350	Aar96350 RGD cycli
155	28	87.5	452	5	ABE10110	Abb10110 Mouse cas	228	27	84.4	7	AAE96324	Aar96324 RGD cycli
156	28	87.5	452	7	ADP79812	Adp79812 Rat caspa	229	27	84.4	7	AAW66882	Aaw66882 Peptide u
157	28	87.5	452	7	ADG63000	Ades63000 Rat Prote	230	27	84.4	7	AAW66905	Aaw66905 Peptide u
158	28	87.5	457	6	ABU26888	Abu26888 Cellbioh	231	27	84.4	7	AAW96922	Aaw96922 Peptide u
159	28	87.5	466	8	ADN20343	Adn20343 Bacterial	232	27	84.4	7	AAW96946	Aaw96946 Peptide u
160	28	87.5	468	3	ADJ70001	Adj70001 Human hea	233	27	84.4	10	ADN99532	Adn99532 Cdc7-ASK
161	28	87.5	469	3	AAE44453	Aay44453 Human car	234	27	84.4	10	ADN99533	Adn99533 Cdc7-ASK
162	28	87.5	469	7	ADG63083	Ades63083 Human per	235	27	84.4	34	AAE54431	Aae54431 Human pan
163	28	87.5	469	7	ADJ69540	Adj69540 Human hea	236	27	84.4	54	AAE64796	Aay64796 Human S
164	28	87.5	469	7	ADJ70494	Adj70494 Human hea	237	27	84.4	68	AAU40615	Aau40615 Propionib
165	28	87.5	469	8	ABM82375	Abm82375 Tumour-as	238	27	84.4	68	ABM37134	Abm37134 Propionib
166	28	87.5	470	8	ADN20289	Adn20289 Bacterial	239	27	84.4	78	AAU91088	Aau91088 Human imm
167	28	87.5	477	7	ABW02002	Abw02002 Human cro	240	27	84.4	79	AU45268	Au45268 Propionib
168	28	87.5	532	8	ADR88978	A. thalia	241	27	84.4	79	ABM41787	Abm41787 Propionib
169	28	87.5	533	8	ADR88968	A. thalia	242	27	84.4	87	ABG16248	Abg16248 Novel hum
170	28	87.5	535	8	ADR88980	A. thalia	243	27	84.4	110	ABE64832	AbE64832 Drosophi1
171	28	87.5	550	7	ADH10080	Adh10080 Human SEA	244	27	84.4	122	ABU34344	Abu34344 Protein e

245	27	84.4	156	4	AAU50461	Aau50461	Propionib	318	27	84.4	603	2	AAU50461	WD-40 dom
246	27	84.4	156	4	ABG28979	Abg28979	Novel hum	319	27	84.4	603	7	ADD48824	Rat Prote
247	27	84.4	156	6	ABM46980	Abm46980	Propionib	320	27	84.4	603	7	ADD48885	Rat Prote
248	27	84.4	164	6	ABU15227	Abu15227	Protein e	321	27	84.4	603	7	ADD48885	Rat Prote
249	27	84.4	165	4	ABU15227	Abu15227	Protein e	322	27	84.4	603	7	ADD48885	Rat Prote
250	27	84.4	166	7	ABU15227	Abu15227	Protein e	323	27	84.4	603	7	ADD48885	Rat Prote
251	27	84.4	170	7	ABU15227	Abu15227	Protein e	324	27	84.4	603	7	ADD48885	Rat Prote
252	27	84.4	182	8	ADN46681	Adn46681	Thermococ	325	27	84.4	605	6	AAO23109	IGFALS 'h
253	27	84.4	190	3	ABG16032	Abg16032	Novel hum	326	27	84.4	605	7	ADD48887	Human Pro
254	27	84.4	196	4	ABG16032	Abg16032	Novel hum	327	27	84.4	605	7	ADD48887	Human Pro
255	27	84.4	214	3	AAU50461	Aau50461	Novel hum	328	27	84.4	605	7	ADD48887	Human Pro
256	27	84.4	214	3	AAU50461	Aau50461	Novel hum	329	27	84.4	605	8	ADJ68826	Human hea
257	27	84.4	214	4	AAU50461	Aau50461	Novel hum	330	27	84.4	605	8	ADJ68826	Human hea
258	27	84.4	214	4	AAU50461	Aau50461	Novel hum	331	27	84.4	605	8	ADJ68826	Human hea
259	27	84.4	214	4	AAU50461	Aau50461	Novel hum	332	27	84.4	605	8	ADJ68826	Human hea
260	27	84.4	214	6	ABU03482	Abu03482	Angiogene	333	27	84.4	605	8	ADJ68826	Human hea
261	27	84.4	214	6	ABU03482	Abu03482	Angiogene	334	27	84.4	605	8	ADJ68826	Human hea
262	27	84.4	219	2	AAU50461	Aau50461	Novel hum	335	27	84.4	605	8	ADJ68826	Human hea
263	27	84.4	219	2	AAU50461	Aau50461	Novel hum	336	27	84.4	605	8	ADJ68826	Human hea
264	27	84.4	220	5	AAU50461	Aau50461	Novel hum	337	27	84.4	605	8	ADJ68826	Human hea
265	27	84.4	225	4	ABU17537	Abu17537	Novel sig	338	27	84.4	605	8	ADJ68826	Human hea
266	27	84.4	225	4	ABU17537	Abu17537	Novel sig	339	27	84.4	605	8	ADJ68826	Human hea
267	27	84.4	225	4	ABU17537	Abu17537	Novel sig	340	27	84.4	605	8	ADJ68826	Human hea
268	27	84.4	225	7	ADU9582	Adu9582	Human pro	341	27	84.4	605	8	ADJ68826	Human hea
269	27	84.4	230	7	ADU9582	Adu9582	Human pro	342	27	84.4	605	8	ADJ68826	Human hea
270	27	84.4	234	8	ADU9582	Adu9582	Human pro	343	27	84.4	605	8	ADJ68826	Human hea
271	27	84.4	240	3	AAU52199	Aau52199	Human tum	344	27	84.4	605	8	ADJ68826	Human hea
272	27	84.4	247	8	ADQ39929	Adq39929	Human myo	345	27	84.4	605	8	ADJ68826	Human hea
273	27	84.4	247	8	ADQ39929	Adq39929	Human myo	346	27	84.4	605	8	ADJ68826	Human hea
274	27	84.4	263	3	AAU00995	Aau00995	Interleuk	347	27	84.4	605	8	ADJ68826	Human hea
275	27	84.4	263	8	ADQ26082	Adq26082	Caspase 1	348	27	84.4	605	8	ADJ68826	Human hea
276	27	84.4	263	8	ADQ26082	Adq26082	Caspase 1	349	27	84.4	605	8	ADJ68826	Human hea
277	27	84.4	280	2	AAU29606	Aau29606	Human ing	350	27	84.4	605	8	ADJ68826	Human hea
278	27	84.4	280	4	AAU29606	Aau29606	Human ing	351	27	84.4	605	8	ADJ68826	Human hea
279	27	84.4	280	8	ADH09685	Adh09685	Human hos	352	27	84.4	605	8	ADJ68826	Human hea
280	27	84.4	280	8	ADH09685	Adh09685	Human hos	353	27	84.4	605	8	ADJ68826	Human hea
281	27	84.4	280	8	ADH09685	Adh09685	Human hos	354	27	84.4	605	8	ADJ68826	Human hea
282	27	84.4	280	8	ADH09685	Adh09685	Human hos	355	27	84.4	605	8	ADJ68826	Human hea
283	27	84.4	280	8	ADH09685	Adh09685	Human hos	356	27	84.4	605	8	ADJ68826	Human hea
284	27	84.4	280	8	ADH09685	Adh09685	Human hos	357	27	84.4	605	8	ADJ68826	Human hea
285	27	84.4	280	8	ADH09685	Adh09685	Human hos	358	27	84.4	605	8	ADJ68826	Human hea
286	27	84.4	283	4	ABG68798	Abg68798	Drosophil	359	27	84.4	605	8	ADJ68826	Human hea
287	27	84.4	286	8	ADQ76230	Adq76230	Human cel	360	27	84.4	605	8	ADJ68826	Human hea
288	27	84.4	291	8	ADQ76230	Adq76230	Human cel	361	27	84.4	605	8	ADJ68826	Human hea
289	27	84.4	291	8	ADQ76230	Adq76230	Human cel	362	27	84.4	605	8	ADJ68826	Human hea
290	27	84.4	299	4	AAU34612	Aau34612	E. coli c	363	27	84.4	605	8	ADJ68826	Human hea
291	27	84.4	299	6	ABU28671	Abu28671	Protein e	364	27	84.4	605	8	ADJ68826	Human hea
292	27	84.4	326	2	AAU28671	Aau28671	Protein e	365	27	84.4	605	8	ADJ68826	Human hea
293	27	84.4	326	2	AAU28671	Aau28671	Protein e	366	27	84.4	605	8	ADJ68826	Human hea
294	27	84.4	333	2	AAU28671	Aau28671	Protein e	367	27	84.4	605	8	ADJ68826	Human hea
295	27	84.4	341	6	ABU34191	Abu34191	Protein e	368	27	84.4	605	8	ADJ68826	Human hea
296	27	84.4	343	6	ABU36699	Abu36699	Bacterial	369	27	84.4	605	8	ADJ68826	Human hea
297	27	84.4	351	7	ADU08053	Adu08053	Bacterial	370	27	84.4	605	8	ADJ68826	Human hea
298	27	84.4	352	3	ABU33165	Abu33165	Macaca mu	371	27	84.4	605	8	ADJ68826	Human hea
299	27	84.4	352	3	ABU33165	Abu33165	Macaca mu	372	27	84.4	605	8	ADJ68826	Human hea
300	27	84.4	395	5	AAU48926	Aau48926	Regulator	373	27	84.4	605	8	ADJ68826	Human hea
301	27	84.4	395	5	AAU48926	Aau48926	Regulator	374	27	84.4	605	8	ADJ68826	Human hea
302	27	84.4	395	5	AAU48926	Aau48926	Regulator	375	27	84.4	605	8	ADJ68826	Human hea
303	27	84.4	395	6	ADU15041	Adu15041	Mouse HYP	376	27	84.4	605	8	ADJ68826	Human hea
304	27	84.4	395	7	ADU15041	Adu15041	Mouse HYP	377	27	84.4	605	8	ADJ68826	Human hea
305	27	84.4	502	4	ABU70242	Abu70242	Drosophil	378	27	84.4	605	8	ADJ68826	Human hea
306	27	84.4	522	4	ABU70242	Abu70242	Drosophil	379	27	84.4	605	8	ADJ68826	Human hea
307	27	84.4	526	6	ABU26885	Abu26885	Cellobioh	380	27	84.4	605	8	ADJ68826	Human hea
308	27	84.4	530	4	ADU89648	Adu89648	Drosophil	381	27	84.4	605	8	ADJ68826	Human hea
309	27	84.4	540	8	ADU89648	Adu89648	Drosophil	382	27	84.4	605	8	ADJ68826	Human hea
310	27	84.4	543	7	ADU89648	Adu89648	Drosophil	383	27	84.4	605	8	ADJ68826	Human hea
311	27	84.4	553	7	ADU89648	Adu89648	Drosophil	384	27	84.4	605	8	ADJ68826	Human hea
312	27	84.4	554	4	ABU57866	Abu57866	Drosophil	385	27	84.4	605	8	ADJ68826	Human hea
313	27	84.4	557	3	AAU87913	Aau87913	D. opposi	386	27	84.4	605	8	ADJ68826	Human hea
314	27	84.4	559	7	ADU79798	Adu79798	Mouse put	387	27	84.4	605	8	ADJ68826	Human hea
315	27	84.4	572	4	ABU57867	Abu57867	Drosophil	388	27	84.4	605	8	ADJ68826	Human hea
316	27	84.4	575	7	ADU5022	Adu5022	Human pro	389	27	84.4	605	8	ADJ68826	Human hea
317	27	84.4	602	5	AAU47598	Aau47598	Drosophil	390	27	84.4	605	8	ADJ68826	Human hea

391	26	81.2	5	7	ABUG31172	Targeting	464	7	ADJ73093	Integrin
392	26	81.2	5	8	ADJ37781	RGD-conta	465	7	ADJ52728	CHI delet
393	26	81.2	5	8	Adm10985	Ultrason	466	7	ADJ51689	CH1 delet
394	26	81.2	6	2	AAR10406	Fibrinose	467	7	ADR75515	Target no
395	26	81.2	6	2	AAR10406	Fibrinose	468	7	ADR75515	Target no
396	26	81.2	6	2	AAR99222	RGD-type	469	8	AAR21003	Cyclised
397	26	81.2	6	2	AAR76191	Alphav/be	470	8	AAR21021	Cyclised
398	26	81.2	6	2	AAR03508	Alpha(5)-	471	8	AAR21004	Cyclised
399	26	81.2	6	2	AAR03483	Alpha(5)-	472	8	AAR99214	RGD-type
400	26	81.2	6	2	AAR35454	Non-dendr	473	8	AAR99223	RGD-type
401	26	81.2	6	2	AAR35453	Non-dendr	474	8	AAR96333	RGD cycli
402	26	81.2	6	2	AAR48528	Integrin	475	8	AAR96340	RGD cycli
403	26	81.2	6	2	AAR48582	Integrin	476	8	AAR96391	RGD cycli
404	26	81.2	6	2	AAR48500	Integrin	477	8	AAR45506	Targeting
405	26	81.2	6	2	AAR21561	Integrin	478	8	AAR45505	Targeting
406	26	81.2	6	2	AAR21586	Integrin	479	8	AAR45507	Targeting
407	26	81.2	6	5	ABB78357	Amino aci	480	8	AAR48471	Integrin
408	26	81.2	6	5	AAM48797	Tumour-ta	481	8	AAR48481	Integrin
409	26	81.2	6	6	ABU59535	Finbronec	482	8	AAR48472	Integrin
410	26	81.2	6	6	ABU59538	Finbronec	483	8	AAR48525	Integrin
411	26	81.2	6	8	ADJ37783	RGD-conta	484	8	AAR48524	Integrin
412	26	81.2	6	8	ADJ37782	RGD-conta	485	8	AAR48526	Integrin
413	26	81.2	7	2	AAR10400	Fibrinose	486	8	AAR48534	Integrin
414	26	81.2	7	2	AAR21010	Cyclised	487	8	AAR48521	Integrin
415	26	81.2	7	2	AAR21006	Cyclised	488	8	AAR66939	Peptide u
416	26	81.2	7	2	AAR57220	Cyclised	489	8	AAR66836	Peptide u
417	26	81.2	7	2	AAR99213	RGD-type	490	8	AAR66911	Peptide u
418	26	81.2	7	2	AAR99219	RGD-type	491	8	ADH67970	Compositi
419	26	81.2	7	2	AAR99220	RGD-type	492	8	ADH67971	Compositi
420	26	81.2	7	2	AAR99220	RGD-type	493	8	ADH67972	Compositi
421	26	81.2	7	2	AAR99208	RGD-type	494	8	AAR96981	Peptide u
422	26	81.2	7	2	AAR99212	RGD-type	495	8	AAR97017	Peptide u
423	26	81.2	7	2	AAR79086	Alpha5/be	496	8	AAR96941	Peptide u
424	26	81.2	7	2	AAR76197	Alpha5/be	497	8	AAR95405	Integrin
425	26	81.2	7	2	AAR96316	RGD cycli	498	8	AAB20604	Targeting
426	26	81.2	7	2	AAR96314	RGD cycli	499	8	AAB20602	Targeting
427	26	81.2	7	2	AAR96323	RGD cycli	500	8	AAB20603	Targeting
428	26	81.2	7	2	AAR96378	RGD cycli	501	8	AAE17094	Cyclic in
429	26	81.2	7	2	AAR96398	RGD cycli	502	8	AAR74977	Transfct
430	26	81.2	7	2	AAR96337	RGD cycli	503	8	AAR74955	Integrin
431	26	81.2	7	2	AAR96339	RGD cycli	504	8	ABG30399	Glycoprot
432	26	81.2	7	2	AAR96339	RGD cycli	505	8	ABG30401	Glycoprot
433	26	81.2	7	2	AAR48532	Integrin	506	8	ABG30400	Glycoprot
434	26	81.2	7	2	AAR48516	Integrin	507	8	ABP98675	Cyclic RG
435	26	81.2	7	2	AAR48556	Integrin	508	8	ABU62647	RGD pepit
436	26	81.2	7	2	AAR48544	Integrin	509	8	ABU63167	Targeting
437	26	81.2	7	2	AAR48473	Integrin	510	8	ABU63168	Targeting
438	26	81.2	7	2	AAR48501	Integrin	511	8	ABU63169	Targeting
439	26	81.2	7	2	AAR48519	Integrin	512	8	ADF82864	Integrin
440	26	81.2	7	2	AAR48531	Integrin	513	8	ADM10980	Ultrason
441	26	81.2	7	2	AAR48555	Integrin	514	8	ADM10981	Ultrason
442	26	81.2	7	2	AAR48522	Integrin	515	8	ADM10982	Ultrason
443	26	81.2	7	2	AAR48596	Integrin	516	8	ADM96259	Human alp
444	26	81.2	7	2	AAR48529	Integrin	517	8	ADS13912	Gene/anti
445	26	81.2	7	2	AAR48523	Integrin	518	9	AAR12332	Platelet-
446	26	81.2	7	2	AAR66936	Peptide u	519	9	AAR12333	Platelet-
447	26	81.2	7	2	AAR66901	Peptide u	520	9	AAR21019	Cyclised
448	26	81.2	7	2	AAR66922	Peptide u	521	9	AAR21009	Cyclised
449	26	81.2	7	2	AAR66859	Peptide u	522	9	AAR21020	Cyclised
450	26	81.2	7	2	AAR66902	Peptide u	523	9	AAR29262	Selectin-
451	26	81.2	7	2	AAR66902	Peptide u	524	9	AAR79074	Alphav/be
452	26	81.2	7	2	AAR66892	Peptide u	525	9	AAR79074	Alphav/be
453	26	81.2	7	2	AAR66921	Peptide u	526	9	AAR79073	Alphav/be
454	26	81.2	7	2	AAR43232	RGD-conta	527	9	AAR76199	Alphav/be
455	26	81.2	7	2	AAR96912	Peptide u	528	9	AAR79089	Alphav/be
456	26	81.2	7	2	AAR96889	Peptide u	529	9	AAR79089	Alphav/be
457	26	81.2	7	2	AAR96921	Peptide u	530	9	AAR79089	Alphav/be
458	26	81.2	7	2	AAR96914	Peptide u	531	9	AAR79089	Alphav/be
459	26	81.2	7	2	AAR96931	Peptide u	532	9	AAR79089	Alphav/be
460	26	81.2	7	3	AAAB17340	Integrin	533	9	AAR79089	Alphav/be
461	26	81.2	7	3	AAAY90212	Alphav in	534	9	AAR79089	Alphav/be
462	26	81.2	7	3	AAAY90219	UPAR targ	535	9	AAR79089	Alphav/be
463	26	81.2	7	5	ABB72939	Integrin	536	9	AAR79089	Alphav/be

537	26	81.2	9	2	AAW48476	Integrin	610	26	81.2	9	7	ADC44955	Endotheli
538	26	81.2	9	2	AAW48477	Integrin	611	26	81.2	9	7	ADF12488	RGD pepti
539	26	81.2	9	2	AAW51235	Cyclic pe	612	26	81.2	9	7	ADF55508	RGD-4C pe
540	26	81.2	9	2	AAW56034	Chimeric	613	26	81.2	9	7	ADJ73101	Integrin
541	26	81.2	9	2	AAW56035	Chimeric	614	26	81.2	9	7	ADJ73100	Integrin
542	26	81.2	9	2	AAW66907	Peptide u	615	26	81.2	9	7	ADJ73099	Integrin
543	26	81.2	9	2	AAW42255	Synthetic	616	26	81.2	9	7	ADJ73115	Integrin
544	26	81.2	9	2	AAW43233	RGD-conta	617	26	81.2	9	7	ADJ73098	Integrin
545	26	81.2	9	2	AAW96977	Peptide u	618	26	81.2	9	8	ADJ52733	CHI delet
546	26	81.2	9	2	AAW93626	NGR recep	619	26	81.2	9	8	ADJ52750	CHI delet
547	26	81.2	9	3	AAW48821	Membrane	620	26	81.2	9	8	ADJ52735	CHI delet
548	26	81.2	9	3	AAW54271	Alpha Vbe	621	26	81.2	9	8	ADJ52734	CHI delet
549	26	81.2	9	3	AAW44970	RGD-4C ta	622	26	81.2	9	8	ADJ52736	CHI delet
550	26	81.2	9	3	AAW17928	TFO-mimet	623	26	81.2	9	8	ADJ51697	CHI delet
551	26	81.2	9	3	AAW17345	Integrin-	624	26	81.2	9	8	ADJ51695	CHI delet
552	26	81.2	9	3	AAW17964	Integrin-	625	26	81.2	9	8	ADJ51694	CHI delet
553	26	81.2	9	3	AAW17346	Integrin-	626	26	81.2	9	8	ADJ51696	CHI delet
554	26	81.2	9	3	AAW17348	Integrin-	627	26	81.2	9	8	ADJ51711	CHI delet
555	26	81.2	9	3	AAW17347	Integrin-	628	26	81.2	9	8	ADJ51516	RGD-conta
556	26	81.2	9	3	AAW90211	Alphav in	629	26	81.2	9	8	ADM72073	tumour ho
557	26	81.2	9	3	AAW22875	Angiogeni	630	26	81.2	9	8	ADJ95924	Antibody-
558	26	81.2	9	3	AAW21701	Human bre	631	26	81.2	9	8	ADM96257	Human alp
559	26	81.2	9	4	AAW20271	Peptide t	632	26	81.2	10	2	AAW11746	Cyclic pl
560	26	81.2	9	4	AAW50242	Enhanced	633	26	81.2	10	2	AAW11747	Cyclic pl
561	26	81.2	9	4	AAW97086	Integrin-	634	26	81.2	10	2	AAW10181	RGD-contg
562	26	81.2	9	4	AAW62279	Tumour ho	635	26	81.2	10	2	AAW76189	Integrin
563	26	81.2	9	4	AAW08989	ICE pepti	636	26	81.2	10	2	AAW98935	Integrin
564	26	81.2	9	4	AAW11044	RGD-conta	637	26	81.2	10	2	AAW03419	Fibronect
565	26	81.2	9	5	AAW98837	Tumour ho	638	26	81.2	10	2	AAW03419	Fibronect
566	26	81.2	9	5	ABG35079	RGD-4C-be	639	26	81.2	10	2	AAW79650	N-alpha-A
567	26	81.2	9	5	AAW81135	Integrin-	640	26	81.2	10	2	AAW79652	N-alpha-A
568	26	81.2	9	5	AAW81110	Integrin-	641	26	81.2	10	2	AAW79651	N-alpha-A
569	26	81.2	9	5	AAW81094	Integrin-	642	26	81.2	10	2	AAW12863	RGD-bindi
570	26	81.2	9	5	AAW81134	Integrin-	643	26	81.2	10	2	AAW35451	Non-dendr
571	26	81.2	9	5	AAW81086	RGD/NGR d	644	26	81.2	10	2	AAW48516	Integrin
572	26	81.2	9	5	ABP54051	Alpha-v-b	645	26	81.2	10	2	AAW48515	Integrin
573	26	81.2	9	5	ABW72944	Integrin	646	26	81.2	10	2	AAW99008	Peptide R
574	26	81.2	9	5	ABW72945	Integrin	647	26	81.2	10	2	AAW21563	Integrin-
575	26	81.2	9	5	ABW72946	Integrin	648	26	81.2	10	2	AAW95406	Integrin-
576	26	81.2	9	5	ABW72947	Integrin	649	26	81.2	10	3	AAW90755	Alphasbet
577	26	81.2	9	5	ABW72961	Integrin	650	26	81.2	10	3	AAW21716	Human tum
578	26	81.2	9	5	ABJ04333	Stem cell	651	26	81.2	10	4	AAW08561	RGD-4C pe
579	26	81.2	9	5	ABJ04322	Bone marr	652	26	81.2	10	5	ABG68408	Transloca
580	26	81.2	9	5	ABJ04451	Stem cell	653	26	81.2	10	5	ABG68408	Transloca
581	26	81.2	9	5	ABJ04359	BRAS11 me	654	26	81.2	10	5	ABG68408	Transloca
582	26	81.2	9	5	ABJ04427	Stem cell	655	26	81.2	10	5	ABG70755	Targeting
583	26	81.2	9	5	ABJ04569	Bone marr	656	26	81.2	10	5	AAE17113	Cyclic in
584	26	81.2	9	5	ABJ01278	Human cas	657	26	81.2	10	5	AAE17110	Cyclic in
585	26	81.2	9	5	ABW08066	Cyclic RG	658	26	81.2	10	5	AAE17095	Cyclic in
586	26	81.2	9	5	ABG70729	avB3 bind	659	26	81.2	10	5	AAE17097	Cyclic in
587	26	81.2	9	5	ABW76442	RGD-4C pe	660	26	81.2	10	5	AAU74988	RGD conta
588	26	81.2	9	5	AAE17983	Human lig	661	26	81.2	10	5	AAU74979	Transfect
589	26	81.2	9	5	ABW78354	Amino aci	662	26	81.2	10	5	AAU74983	Transfect
590	26	81.2	9	5	ABG70754	Targeting	663	26	81.2	10	6	AAE35396	Peptide c
591	26	81.2	9	5	AAE17109	Cyclic in	664	26	81.2	10	6	AAE35396	Peptide c
592	26	81.2	9	5	AAE17108	Cyclic in	665	26	81.2	10	7	AAO23409	West nile
593	26	81.2	9	5	AAU75609	Synthetic	666	26	81.2	10	7	ABR56197	RGD pepti
594	26	81.2	9	5	AAU79138	Synthetic	667	26	81.2	10	7	ADC33483	Synthetic
595	26	81.2	9	5	AAU98972	Adeno-agg	668	26	81.2	10	8	ADJ82865	Integrin-
596	26	81.2	9	5	AAU74978	Transfect	669	26	81.2	10	8	ADJ37784	RGD-conta
597	26	81.2	9	5	AAW48795	Tumour-ta	670	26	81.2	10	8	ADJ37784	RGD-conta
598	26	81.2	9	5	ABW79525	RGD motif	671	26	81.2	10	8	ADR75421	Target mo
599	26	81.2	9	5	ABG31063	Alpha v b	672	26	81.2	10	8	ADR75430	Target mo
600	26	81.2	9	5	AAW78427	Cyclic pe	673	26	81.2	10	8	ADJ17359	Peptide a
601	26	81.2	9	5	AAW51995	Drug targ	674	26	81.2	10	8	ADJ13913	Gene/anti
602	26	81.2	9	5	AAW51996	Integrin	675	26	81.2	11	2	AAW76194	Integrin
603	26	81.2	9	5	AAW51997	Integrin	676	26	81.2	11	2	AAW11184	Free pept
604	26	81.2	9	6	ABG73024	Integrin	677	26	81.2	11	2	AAW60299	tumour ho
605	26	81.2	9	6	ABW84641	Human int	678	26	81.2	11	2	AAW57199	RGD-conta
606	26	81.2	9	6	ABW59556	Tumour an	679	26	81.2	11	2	AAW95408	Integrin-
607	26	81.2	9	6	ABU08571	RGD-conta	680	26	81.2	11	2	AAW95409	Integrin-
608	26	81.2	9	6	ABP60343	Tumour va	681	26	81.2	11	2	AAW95407	Integrin-
609	26	81.2	9	6	ADA20234	Synthetic	682	26	81.2	11	3	AAW54273	Peptide 1

683	26	81.2	11	3	AAY90159	Aay90159 UPAR targ	756	19	6	AAE38405	Aae38405 HIV 1 Tat
684	26	81.2	11	3	AAV58860	Aay58860 Membrane	757	26	3	ADF77514	DAF-7 pro
685	26	81.2	11	4	AAE06294	Aae06294 Double cy	758	20	2	ABO6057	Caenorhab
686	26	81.2	11	5	AAU97578	Aau97578 Synthetic	759	26	4	AAm51432	Peptide K
687	26	81.2	11	5	AAU97580	Aau97580 Synthetic	760	20	4	AAE38386	Erbb2 bin
688	26	81.2	11	5	AAU97577	Aau97577 Synthetic	761	20	6	ADC33480	Peptide K
689	26	81.2	11	5	AAU97579	Aau97579 Synthetic	762	20	7	ADE86004	Human Hu-
690	26	81.2	11	5	AAU97581	Aau97581 Synthetic	763	26	8	ADQ76841	Mouse alp
691	26	81.2	11	5	AAU87024	Aau87024 Targettin	764	26	8	ADQ76841	Mouse alp
692	26	81.2	11	5	AAE17096	Aae17096 Cyclic in	765	26	21	AAW96218	AlphavBet
693	26	81.2	11	5	AAE17098	Aae17098 Cyclic in	766	26	21	ADE03400	GPIIb/III
694	26	81.2	11	5	AAO21743	Aao21743 Procytoto	767	26	21	ADN11199	Peptide m
695	26	81.2	11	5	AAE33595	Aae33595 Peptide c	768	26	21	ADN11233	Peptide m
696	26	81.2	11	6	ABP58062	Abp58062 RGD pepti	769	26	8	ADN11234	Peptide m
697	26	81.2	11	6	ABP70218	Abp70218 Amino aci	770	26	21	ADN11200	Peptide m
698	26	81.2	11	6	ABR40189	Abra40189 Integrin	771	26	8	ADM96250	Human cas
699	26	81.2	11	7	ADF12487	Adf12487 RGD pepti	772	26	8	ADR41069	Cyclic KD
700	26	81.2	11	8	ADF82866	Adf82866 Integrin-	773	26	8	ADR51117	Duo bindi
701	26	81.2	11	8	ADF82867	Adf82867 Integrin-	774	26	22	AAm15231	Peptide #
702	26	81.2	11	8	ADF82868	Adf82868 Integrin-	775	26	22	ABb34221	Peptide #
703	26	81.2	11	8	ADN11238	Adn11238 Peptide u	776	26	22	AAm27692	Peptide #
704	26	81.2	11	8	ADO80846	Ado80846 Synthetic	777	26	22	ABB29060	Peptide #
705	26	81.2	11	8	ADR12199	Adr12199 RGD-4C pe	778	26	22	ABB19657	Protein #
706	26	81.2	11	8	ADS13914	Ads13914 Gene/anti	779	26	22	AAm67400	Human bon
707	26	81.2	11	8	ADS13915	Ads13915 Gene/anti	780	26	22	AAm55015	Human bra
708	26	81.2	11	8	ADS13916	Ads13916 Gene/anti	781	26	22	ABG49059	Human liv
709	26	81.2	12	2	AAW56052	Aaw56052 Chimeric	782	26	22	AAm02974	Peptide #
710	26	81.2	12	2	AAW50655	Aaw50655 Integrin	783	26	22	AAW56044	Chimeric
711	26	81.2	12	2	AAW95410	Aaw95410 Integrin-	784	26	22	AAW96220	Modified
712	26	81.2	12	4	AAW51437	Aam51437 Carboxyl	785	26	23	AAW56044	Chimeric
713	26	81.2	12	5	ABP56215	Abp56215 Chimeric	786	26	24	AAW56044	Chimeric
714	26	81.2	12	5	ABP56186	Abp56186 Chimeric	787	26	24	AAW95097	RGD/SV40
715	26	81.2	12	5	AAE17099	Aae17099 Cyclic in	788	26	24	AAV79875	Integrin-
716	26	81.2	12	5	AAE17103	Aae17103 Cyclic in	789	26	24	AAV83259	Integrin
717	26	81.2	12	5	AAU74965	Aau74965 RGD conta	790	26	24	AAV96572	Integrin-
718	26	81.2	12	8	ADF82869	Adf82869 Integrin-	791	26	24	AAE60908	RGD pepri
719	26	81.2	12	8	ADF82876	Adf82876 Integrin-	792	26	24	ABE05522	Integrin-
720	26	81.2	12	8	ADF82877	Adf82877 Integrin-	793	26	24	ABE05522	Integrin-
721	26	81.2	12	8	ADS13917	Ads13917 Gene/anti	794	26	24	ABP56201	Chimeric
722	26	81.2	13	2	AAV43226	Aay43226 RGD-Conta	795	26	24	ABP56184	Chimeric
723	26	81.2	13	2	AAV43229	Aay43229 RGD-Conta	796	26	24	ABP56187	Chimeric
724	26	81.2	13	2	AAV48725	Aay48725 Membrane	797	26	24	ABP56202	Chimeric
725	26	81.2	13	3	AAV90168	Aay90168 UPAR targ	798	26	24	ADF70993	Phosphope
726	26	81.2	13	3	AAV90158	Aay90158 UPAR targ	799	26	24	ADM82999	Integrin-
727	26	81.2	13	5	AAU98801	Aau98801 Peptide 1	800	26	25	AAE21940	Homing an
728	26	81.2	13	5	AAU98802	Aau98802 Peptide 1	801	26	25	AAE06517	Homing pr
729	26	81.2	13	6	ABU63059	Abu63059 Peptide 1	802	26	25	ABP56213	Chimeric
730	26	81.2	13	6	ABU63059	Abu63059 Peptide 1	803	26	25	ABP56212	Chimeric
731	26	81.2	14	2	AAW19833	Aaw19833 RGD pepti	804	26	25	ABP56211	Chimeric
732	26	81.2	14	2	AAW56051	Aaw56051 Chimeric	805	26	25	ABP56214	Chimeric
733	26	81.2	14	2	AAW50652	Aaw50652 DAB conta	806	26	25	ABP56185	Chimeric
734	26	81.2	14	3	AAV59625	Aay59625 Maturatio	807	26	26	AAE21937	Homing an
735	26	81.2	14	3	AAV98966	Aay98966 Peptide #	808	26	26	AAE06516	Homing pr
736	26	81.2	14	8	ADM96254	Adm96254 Apoptosis	809	26	26	AAE17092	Cyclic in
737	26	81.2	15	2	AAE62581	Aae62581 Human hep	810	26	26	AAE17092	Cyclic in
738	26	81.2	15	2	AAW56040	Aaw56040 Chimeric	811	26	26	AAU74953	RGD conta
739	26	81.2	15	2	AAV43228	Aay43228 RGD-Conta	812	26	26	AAU48798	Tumour-ta
740	26	81.2	15	2	AAV43228	Aay43228 RGD-Conta	813	26	26	ADS13910	Gene/anti
741	26	81.2	15	3	AAV54272	Aay54272 Peptide i	814	26	26	AAW95402	Conserved
742	26	81.2	15	3	AAV90167	Aay90167 UPAR targ	815	26	27	AAW95404	Conserved
743	26	81.2	15	6	ABU59649	Abu59649 PEG conta	816	26	27	AAE47345	DNA bindi
744	26	81.2	15	8	AAO24331	Aao24331 Human Sli	817	26	27	AAE06191	Human gen
745	26	81.2	16	7	ADR03397	Ado03397 GPIIb/III	818	26	27	ABG34015	Human sec
746	26	81.2	16	7	ADF12496	Adf12496 Mouse p35	819	26	27	AAE17120	Integrin-
747	26	81.2	17	2	AAV20690	Aay20690 Human neu	820	26	27	AAE17093	Cyclic in
748	26	81.2	18	2	AAW79487	Aaw79487 Peptide i	821	26	27	AAE17122	Peptide x
749	26	81.2	18	3	AAE49453	Aae49453 Peptide s	822	26	27	AAE17091	Cyclic in
750	26	81.2	18	4	AAE04343	Aae04343 Human dis	823	26	27	AAE17125	Integrin-
751	26	81.2	18	5	AAE23768	Aae23768 Zdscl p3	824	26	27	AAU74970	Alpha V,
752	26	81.2	18	7	ADF12495	Adf12495 Canine p3	825	26	27	AAU74952	RGD conta
753	26	81.2	18	7	ADF12494	Adf12494 Human p35	826	26	27	AAU74974	Integrin
754	26	81.2	19	5	AAm51998	Aam51998 Integrin	827	26	27	AAU74954	RGD conta
755	26	81.2	19	6	AAE38404	Aae38404 HIV 1 Tat	828	26	27	ABE08603	Gene tran
								27	6	ABE98397	Integrin

829	26	81.2	27	8	ADS13911	Adsl3911 Gene/anti	902	26	81.2	60	4	AAU61298	Aau61298 Propionib
830	26	81.2	27	8	ADS13909	Adsl3909 Gene/anti	903	26	81.2	60	5	ABP10154	Abp10154 Human ORF
831	26	81.2	28	2	AAW95401	Aaw95401 Conserved	904	26	81.2	60	5	ABM57817	Abm57817 Propionib
832	26	81.2	28	3	AAW95401	Aaw95401 Conserved	904	26	81.2	60	5	ABM57817	Abm57817 Propionib
833	26	81.2	28	3	AAW95401	Aaw95401 Conserved	905	26	81.2	60	6	ABM38744	Abm38744 Human alp
834	26	81.2	28	5	AAE17123	Aae17123 Integrin-	906	26	81.2	60	8	ABM96252	Abm96252 Human alp
835	26	81.2	28	5	AAE17090	Aae17090 Cyclic RG	907	26	81.2	62	4	ABG07543	Abg07543 Novel hum
836	26	81.2	28	5	AAU74973	Aau74973 Alpha V i	908	26	81.2	62	4	ABG13743	Abg13743 Novel hum
837	26	81.2	28	5	AAU74951	Aau74951 RGD conta	909	26	81.2	63	3	AAI49443	Aay49443 Murine Zd
838	26	81.2	28	8	ADN11184	Adn11184 Peptide i	910	26	81.2	63	3	AAI49443	Aay49443 Murine Zd
839	26	81.2	28	8	ADS13908	Adsl3908 Gene/anti	911	26	81.2	63	3	AAI49443	Aay49443 Murine Zd
840	26	81.2	29	4	AAE02759	Aae02759 Kalata B1	912	26	81.2	63	4	AAU58320	Aau58320 Propionib
841	26	81.2	29	5	ABU67204	Abu67204 G-protein	913	26	81.2	63	4	AAU58320	Aau58320 Propionib
842	26	81.2	29	6	ABP53985	Abp53985 Human adr	914	26	81.2	63	4	AAU58320	Aau58320 Propionib
843	26	81.2	29	8	ADO28725	Ado28725 Human alp	915	26	81.2	63	5	AAE23758	Aae23758 Murine T17
844	26	81.2	29	8	ADO05229	Ado05229 Adrenergi	916	26	81.2	63	5	AAE23758	Aae23758 Murine T17
845	26	81.2	30	2	AAW48646	Aaw48646 N-termina	917	26	81.2	63	5	AAE21091	Aae21091 Mouse T17
846	26	81.2	30	3	AAW48646	Aaw48646 Lactoferr	918	26	81.2	63	6	ABM54839	Abm54839 Propionib
847	26	81.2	35	2	ADP77517	Adp77517 DAP-7 pro	919	26	81.2	64	4	AAU40465	Aau40465 Propionib
848	26	81.2	35	3	ABO60600	Ab060600 Caenorhab	920	26	81.2	64	4	AAU40465	Aau40465 Propionib
849	26	81.2	39	3	AAW49444	Aaw49444 Murine Zd	921	26	81.2	65	3	ABM36984	Abm36984 Propionib
850	26	81.2	39	4	AAE04334	Aae04334 Murine ma	922	26	81.2	65	4	AAU62810	Aau62810 Propionib
851	26	81.2	39	5	AAE23759	Aae23759 Murine Zd	923	26	81.2	65	5	ABP35285	Abp35285 Human ORF
852	26	81.2	39	7	AAE21094	Aae21094 Mouse T17	924	26	81.2	65	6	ABM59329	Abm59329 Propionib
853	26	81.2	39	8	ABO55221	Ab055221 Mouse mat	925	26	81.2	65	7	ADD89952	Add89952 Foot-and-
854	26	81.2	43	2	ADP77515	Adp77515 DAP-7 pro	926	26	81.2	65	8	ADJ56911	Adj56911 Foot and
855	26	81.2	43	3	ABO60508	Ab060508 Caenorhab	927	26	81.2	65	8	ABO56543	Abo56543 Human gen
856	26	81.2	43	4	AAW49744	Aaw49744 Antimicro	928	26	81.2	67	4	AAU63498	Aau63498 Propionib
857	26	81.2	45	2	AAW05875	Aaw05875 Intermedi	929	26	81.2	67	6	ABM60017	Abm60017 Propionib
858	26	81.2	45	4	ABG25537	Abg25537 Novel hum	930	26	81.2	68	5	ABP09222	Abp09222 Human ORF
859	26	81.2	45	7	ADP69847	Adp69847 AcMA-type	931	26	81.2	68	5	ABP09222	Abp09222 Human ORF
860	26	81.2	46	4	AAE08946	Aae08946 ICE pepti	932	26	81.2	69	3	ABM25079	Abm25079 Plant SDF
861	26	81.2	46	4	AAE08946	Aae08946 ICE pepti	933	26	81.2	69	4	AAU52051	Aau52051 Propionib
862	26	81.2	46	8	ADP52031	Adp52031 Aspartate	934	26	81.2	69	4	AAU44985	Aau44985 Propionib
863	26	81.2	46	8	ADP52031	Adp52031 Aspartate	934	26	81.2	69	4	AAU44985	Aau44985 Propionib
864	26	81.2	48	4	ABG18211	Abg18211 Novel hum	935	26	81.2	69	6	ABM46293	Abm46293 Propionib
865	26	81.2	48	2	AAW02648	Aaw02648 Mutant di	936	26	81.2	69	6	ABM48570	Abm48570 Propionib
866	26	81.2	50	4	AAW80383	Aaw80383 Human hae	937	26	81.2	69	6	ABM41504	Abm41504 Propionib
867	26	81.2	50	4	AAU50635	Aau50635 Propionib	938	26	81.2	70	4	AAW82385	Aaw82385 Human imm
868	26	81.2	50	4	AAU45429	Aau45429 Propionib	939	26	81.2	70	4	AAU53860	Aau53860 Propionib
869	26	81.2	50	6	ABM47154	Abm47154 Propionib	940	26	81.2	70	6	ABM50379	Abm50379 Propionib
870	26	81.2	52	4	AAU46051	Aau46051 Propionib	941	26	81.2	72	2	AAU49774	Aau49774 Propionib
871	26	81.2	52	4	AAU46051	Aau46051 Propionib	942	26	81.2	72	2	AAU49774	Aau49774 Propionib
872	26	81.2	52	4	AAU42570	Abm42570 Propionib	943	26	81.2	72	2	AAU42321	Aau42321 Hirudin (
873	26	81.2	53	6	ABM459523	Abm459523 Propionib	944	26	81.2	72	2	AAU42321	Aau42321 Hirudin (
874	26	81.2	53	6	ABM459523	Abm459523 Propionib	945	26	81.2	72	2	AAU42321	Aau42321 Hirudin (
875	26	81.2	54	4	AAU59762	Aau59762 Propionib	946	26	81.2	76	3	AAU50379	Aau50379 Propionib
876	26	81.2	54	6	ABM56281	Abm56281 Propionib	947	26	81.2	76	3	AAU50379	Aau50379 Propionib
877	26	81.2	55	4	AAU42049	Aau42049 Propionib	948	26	81.2	76	4	ABG02905	Abg02905 Novel hum
878	26	81.2	55	4	AAU39246	Aau39246 Propionib	949	26	81.2	76	4	ABG02905	Abg02905 Novel hum
879	26	81.2	55	4	AAU64092	Aau64092 Propionib	950	26	81.2	76	6	ABM53553	Abm53553 Propionib
880	26	81.2	55	4	AAU56638	Aau56638 Propionib	951	26	81.2	76	7	ABO69836	Abo69836 Pseudomon
881	26	81.2	55	6	ABM60611	Abm60611 Propionib	952	26	81.2	77	2	AAU704834	Aau704834 Mycobacte
882	26	81.2	55	6	ABM35765	Abm35765 Propionib	953	26	81.2	77	4	AAU43934	Aau43934 Propionib
883	26	81.2	55	6	ABM38568	Abm38568 Propionib	954	26	81.2	77	4	AAU43934	Aau43934 Propionib
884	26	81.2	55	6	ABM53157	Abm53157 Propionib	955	26	81.2	77	6	ABM40453	Abm40453 Propionib
885	26	81.2	56	3	AAE25002	Aae25002 Plant SDF	956	26	81.2	77	6	ABM50830	Abm50830 Propionib
886	26	81.2	56	3	AAE25002	Aae25002 Plant SDF	957	26	81.2	79	4	AAU487470	Aau487470 Human imm
887	26	81.2	56	3	AAO09402	Aao09402 Human pol	958	26	81.2	83	5	ABP03980	Abp03980 Human ORF
888	26	81.2	57	3	AAU97990	Aau97990 RGD-proin	959	26	81.2	85	4	ABM88462	Abm88462 Human imm
889	26	81.2	57	8	ABO54720	Abo54720 Human gen	960	26	81.2	85	5	ABP05804	Abp05804 Human ORF
890	26	81.2	58	3	AAE25001	Aae25001 Plant SDF	961	26	81.2	86	3	AAU19125	Aau19125 Zea mays
891	26	81.2	58	4	AAU21186	Aau21186 Peptide #	962	26	81.2	86	4	AAU40276	Aau40276 Human pol
892	26	81.2	58	4	ABM343506	Abm343506 Peptide #	963	26	81.2	86	8	AAU57808	Aau57808 Human gen
893	26	81.2	58	4	AAU37401	Aau37401 Peptide #	964	26	81.2	87	4	AAU42944	Aau42944 Propionib
894	26	81.2	58	4	ABM26466	Abm26466 Protein #	965	26	81.2	87	6	ABM39463	Abm39463 Propionib
895	26	81.2	58	4	AAW64441	Aaw64441 Human bra	966	26	81.2	88	4	AAU20304	Aau20304 Human nov
896	26	81.2	58	4	ABM58884	Abm58884 Human liv	967	26	81.2	89	3	ABM25479	Abm25479 Pinus rad
897	26	81.2	58	5	ABG46265	Abg46265 Human pep	968	26	81.2	89	3	ABM25479	Abm25479 Pinus rad
898	26	81.2	58	5	ABP34415	Abp34415 Human ORF	969	26	81.2	89	3	ABM25479	Abm25479 Pinus rad
899	26	81.2	59	3	AAE25081	Aae25081 Plant SDF	970	26	81.2	90	3	AAU39777	Aau39777 Propionib
900	26	81.2	59	3	AAE25000	Aae25000 Plant SDF	971	26	81.2	90	4	AAU39777	Aau39777 Propionib
901	26	81.2	60	4	AAU42225	Aau42225 Propionib	972	26	81.2	90	6	AAU59831	Aau59831 Propionib
							973	26	81.2	90	6	ABM36296	Abm36296 Propionib
							974	26	81.2	90	6	ABM56350	Abm56350 Propionib

975 26 81.2 91 4 AAU48465
 976 26 81.2 91 6 ABM44984
 977 26 81.2 93 8 ADK19706
 978 26 81.2 94 2 AAY35872
 979 26 81.2 96 4 ABG21787
 980 26 81.2 97 4 AAM06515
 981 26 81.2 97 4 AAM06799
 982 26 81.2 99 2 AAY04833
 983 26 81.2 99 4 AAU40436
 984 26 81.2 99 6 AAW36955
 985 26 81.2 101 4 AAO12673
 986 26 81.2 102 5 ADK36904
 987 26 81.2 103 7 ADB64460
 988 26 81.2 106 4 ABG09018
 989 26 81.2 108 3 AAB53252
 990 26 81.2 108 5 ABP04690
 991 26 81.2 108 8 ADS12498
 992 26 81.2 108 8 ADS12482
 993 26 81.2 110 4 ABG26316
 994 26 81.2 110 7 ADE08262
 995 26 81.2 110 8 ADS12494
 996 26 81.2 111 2 AAW44079
 997 26 81.2 111 2 AAW27643
 998 26 81.2 112 4 AAO00116
 999 26 81.2 115 4 AAO06254
 1000 26 81.2 115 4 AAU48996

AAU48465 Propionib
 ABM44984 Propionib
 ADK19706 Mouse pro
 AAY35872 Chlamydia
 ABG21787 Novel hum
 AAM06515 Human foe
 AAM06799 Human foe
 AAY04833 Mycobacte
 AAU40436 Propionib
 AAW36955 Propionib
 AAO12673 Human pol
 ADK36904 Novel hum
 ADB64460 Human pro
 ABG09018 Novel hum
 AAB53252 Human col
 ABP04690 Human ORF
 ADS12498 Human IGF
 ADS12482 Human IGF
 ABG26316 Novel hum
 ADE08262 Novel pro
 ADS12494 Human IGF
 AAW44079 Human sec
 AAW27643 Secreted
 AAO00116 Human pol
 AAO06254 Human pol
 AAU48996 Propionib

ALIGNMENTS

RESULT 1

AAR96393
 ID AAR96393 standard; peptide; 7 AA.

XX AAR96393;

DT 05-JUL-1996 (first entry)

DE RGD cyclic peptide, TL#86, binds alpha-v,beta3 integrin receptor.

XX RGD-containing peptide; alpha-v, beta-3 integrin receptor; osteoclast;
 matrix; bone; inhibition; bone resorption; promote; endothelial cell;
 smooth muscle cell; restenosis; angiogenesis; cyclic.

XX Synthetic.

XX Key Location/Qualifiers

FT Disulfide-bond 1..7

FT Modified-site 1

FT /note= "Acyl-Cys"

FT Modified-site 7

FT /note= "Amidated C-terminal"

XX WO9528426-A2.

XX 26-OCT-1995.

XX 12-APR-1995; 9SWO-US004741.

XX 13-APR-1994; 94US-00227316.

XX 08-SEP-1994; 94US-00303052.

XX (LJOL-) LA JOLLA CANCER RES FOUND.

XX Cheng S, Ingram R, Mullen D, Tschopp J;

XX WPI; 1995-373767/48.

XX Altering alpha-v, beta-3 integrin receptor-mediated binding of cell to
 matrix - using conformationally restrained peptide of RGD type, e.g. for
 treating inappropriate angiogenesis or for inhibiting bone resorption.

PS Claim 85; Page 79; 99pp; English.

XX The sequences given in AAR96301-417 are non-naturally occurring RGD-
 containing peptides which alter the alpha-v, beta-3 integrin receptor to a
 binding of a cell to a matrix, such as the binding of an osteoclast to a
 matrix such as bone. These peptides inhibit bone resorption and can
 inhibit or promote alpha-v, beta-3-mediated cell attachment depending on
 whether they are present in the cell in a soluble form or are bound to a
 solid substrate. These peptides can be used in the amelioration of the
 severity of a pathology involving alpha-v, beta-3 receptor-mediated
 binding of a cell, such as an osteoclast, endothelial cell or smooth
 muscle cell to a matrix. They are used for treating conditions associated
 with restenosis or inappropriate or insufficient angiogenesis, or for
 inhibiting osteoclast binding to the matrix

SQ Sequence 7 AA;

Query Match 100.0%; Score 32; DB 2; Length 7;

Best Local Similarity 100.0%; Pred. No. 1.8e+06;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5

Db 1 CRGDD 5

RESULT 2

AAW66831

ID AAW66831 standard; peptide; 7 AA.

XX AAW66831;

DT 25-MAR-2003 (revised)

DT 10-DEC-1998 (first entry)

DE Peptide useful for altering bone resorption.

XX bone resorption; pharmacore; angiogenesis; restenosis; integrin receptor;
 alpha v beta 3 integrin receptor; osteoclast; cyclic.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1

FT /note= "N-terminal acetyl"

FT Modified-site 7

FT /note= "C-terminal amide"

XX US5807819-A.

XX 15-SEP-1998.

XX 12-APR-1995; 95US-00421698.

XX 15-APR-1994; 94US-00227316.

XX 08-SEP-1994; 94US-00303052.

XX (LJOL-) LA JOLLA CANCER RES CENT.

XX Tschopp JF, Mullen D, Ingram R, Cheng S;

XX WPI; 1998-555601/47.

XX Use of peptide derivatives which can alter integrin receptor binding -
 for altering bone resorption, treating angiogenesis or restenosis and
 altering integrin receptor mediated interactions.

XX Example 2; Fig 2C; 87pp; English.

XX A new method is claimed for altering bone resorption. It comprises
 administration of a peptide of formula: X1X2X3X4GX5X6X7X8; where X1 =
 R1R2N or 0-10 amino acids (optionally protected by acetylation at the N-
 terminus); X2 = absent or 1 amino acid; X3 = absent or 1 or 2 amino acids

Sequence 9 AA:

Query Match 100.0%; Score 32; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
DB 2 CRGDD 6

RESULT 5
AAW66840
ID AAW66840 standard; peptide; 9 AA.
XX
AC AAW66840;
XX
DT 25-MAR-2003 (revised)
DT 10-DEC-1998 (first entry)
XX
DE Cyclic peptide.
XX
KW bone resorption; pharmacore; angiogenesis; restenosis; integrin receptor;
KW alpha v beta 3 integrin receptor; osteoclast; cyclic.
XX
OS Synthetic.
XX
PN US5807819-A.
XX
PD 15-SEP-1998.
XX
PF 12-APR-1995; 95US-00421698.
XX
PR 15-APR-1994; 94US-00227316.
PR 08-SEP-1994; 94US-00303052.
XX
PA (LJOL-) LA JOLLA CANCER RES CENT.
XX
PI Tschopp JF, Mullen D, Ingram R, Cheng S;
XX
DR WPI; 1998-555601/47.

Use of peptide derivatives which can alter integrin receptor binding -
PT for altering bone resorption, treating angiogenesis or restenosis and
PT altering integrin receptor mediated interactions.

PS Disclosure; Col 47-48; 87pp; English.
XX
XX A new method is claimed for altering bone resorption. It comprises
CC administration of a peptide of formula: X1X2X3X4X5X6X7X8 ; where $\text{X1} =$
CC R1R2N or 0-10 amino acids (optionally protected by acetylation at the N-
CC terminus); $\text{X2} =$ absent or 1 amino acid; $\text{X3} =$ absent or 1 or 2 amino acids
CC; $\text{X4} = \text{N-Me-Arg}$; $\text{X5} =$ residue which provides an ionic interaction with an
CC integrin receptor, or is Mse , Paa or Tfse ; $\text{X6} =$ residue which has an
CC aliphatic side chain; a non-natural amino acid that is hydrophobic; or
CC Thr ; $\text{X7} =$ a residue capable of forming a bond (i) with a bridging amino
CC acid of X2 , (ii) with X3 when X2 is absent, or (iii) with X4 when X2 and
CC X3 are absent, to conformationally restrain the peptide; $\text{X8} = \text{NR3R4}$; OR5;
CC or 0-10 amino acids, optionally protected as an amide at the C-terminus;
CC R1 , $\text{R3-R5} = \text{H}$ or alkyl; $\text{R2} = \text{H}$, alkyl, alkyl-CO or phenyl-CO. The
CC peptides are useful for inhibiting bone resorption, angiogenesis or
CC restenosis, and for altering integrin receptor-mediated interactions,
CC especially alpha v beta 3 integrin receptor-mediated binding of cells to
CC a matrix. They may be used for reducing or inhibiting osteoclast binding
CC to a matrix. The present sequence is shown in the specification. (Updated
CC on 25-MAR-2003 to correct PF field.)
XX

Sequence 9 AA;

Query Match 100.0%; Score 32; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
DB 2 CRGDD 6

Db 2 CRGDD 6
RESULT 6
AAW97003
ID AAW97003 standard; peptide; 9 AA.
XX
AC AAW97003;
XX
DT 26-APR-1999 (first entry)
XX

Peptide used to inhibit bone resorption, angiogenesis and restenosis.

RGD-peptide; bone resorption; angiogenesis; restenosis;

integrin receptor mediated interaction; binding; osteoclast; bone.

XX Synthetic.

XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "acetylated residue"
FT Modified-site 9
FT /note= "amidated residue"

XX US5849865-A.

XX 15-DEC-1998.

XX 12-APR-1995; 95US-00421695.

XX 13-APR-1994; 94US-00227316.

XX 08-SEP-1994; 94US-00303052.

XX (LJOL-) LA JOLLA CANCER RES FOUND.

XX Tschopp JF, Mullen D, Ingram R, Cheng S;
XX
XX WPI; 1999-179486/15.

XX RGD peptides - that modulate integrin-mediated cell-matrix binding.
XX
XX Claim 6; Col 138; 90pp; English.

XX The present peptide represents a non-naturally occurring Arg-Gly-Asp
CC (RGD)-containing peptide. The peptide may be useful for reducing or
CC inhibiting bone resorption, angiogenesis or restenosis, and for altering
CC an integrin receptor mediated interaction. The peptide may also be used
CC to inhibit the binding of an osteoclast to bone
XX
XX Sequence 9 AA;

Query Match 100.0%; Score 32; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
DB 2 CRGDD 6

RESULT 7
AAU61299
ID AAU61299 standard; protein; 92 AA.

XX AAU61299;

XX 27-FEB-2002 (first entry)

XX Propionibacterium acnes immunogenic protein #22195.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;

KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US012865.
 XX
 PR 21-APR-2000; 2000US-0199047P.
 PR 02-JUN-2000; 2000US-0208841P.
 PR 07-JUL-2000; 2000US-0216747P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'Maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 DR WPI; 2001-616774/71.
 DR N-PSDB; AAS59616.
 XX
 XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.
 XX
 PS Example 1; SEQ ID NO 22494; 1069pp; English.
 XX
 XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 92 AA;
 Query Match 100.0%; Score 32; DB 4; Length 92;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CRGDD 5
 Db 75 CRGDD 79
 RESULT 8
 ABM57818
 ID ABM57818 standard; protein; 92 AA.
 XX
 AC ABM57818;
 XX
 XX 20-OCT-2003 (first entry)
 DT
 XX Propionibacterium acnes predicted ORF-encoded polypeptide #22494.
 DE
 XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;
 KW immunostimulant; immune response; vaccine.
 XX

OS Propionibacterium acnes.
 XX
 PN WO2003033515-A1.
 XX
 PD 24-APR-2003.
 XX
 PF 11-OCT-2002; 2002WO-US032727.
 XX
 PR 15-OCT-2001; 2001US-00978825.
 XX
 XX (CORI-) CORIXA CORP.
 PA
 PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
 PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
 PI Barth B, Vallieue-Douglas J;
 XX
 DR WPI; 2003-381789/36.
 DR N-PSDB; ACF64545.
 XX
 XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
 PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
 PT or for stimulating an immune response specific for a P. acnes protein.
 XX
 PS Example 1; SEQ ID NO 22494; 1481pp; English.
 XX
 XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
 CC encoding a Propionibacterium acnes protein. The invention also relates to
 CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
 CC immunogenic fragments of P. acnes polypeptides. The invention
 CC additionally encompasses expression vectors and host cells comprising a
 CC polynucleotide of the invention; antibodies against polypeptides of the
 CC invention; fusion proteins comprising a polypeptide of the invention; a
 CC method for stimulating an immune response specific for a P. acnes
 CC polypeptide and an isolated T cell population comprising T cells prepared
 CC via this method; a vaccine composition (comprising P. acnes polypeptides,
 CC polynucleotides, antibodies, fusion proteins, T cell populations, or
 CC antigen-presenting cells that express the polypeptide); a method and kit
 CC for detecting or determining the presence or absence of P. acnes in a
 CC patient; and a method for inhibiting the development of P. acnes in a
 CC protein. The P. acnes polypeptides, polynucleotides, antibodies, fusion
 CC proteins, T cell populations or antigen-presenting cells that express the
 CC polypeptides are useful for diagnosing, preventing or treating acne
 CC vulgaris, or for stimulating an immune response specific for a P. acnes
 CC protein. The polynucleotides can also be used as probes or primers for
 CC nucleic acid hybridization. The vaccine composition is useful for the
 CC stimulation of an immune response against P. acnes, or for treating acne,
 CC and the kit is useful for performing a diagnostic assay. The present
 CC sequence represents a polypeptide predicted to be encoded by an ORF (open
 CC reading frame) contained within the P. acnes polynucleotides of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 92 AA;
 Query Match 100.0%; Score 32; DB 6; Length 92;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CRGDD 5
 Db 75 CRGDD 79
 RESULT 9
 ABR41549
 ID ABR41549 standard; protein; 117 AA.
 XX
 AC ABR41549;
 XX
 XX 02-JUN-2003 (first entry)
 DT
 XX Human DTHP nucleic acid synthesis/modification protein.
 DE

XX Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis;
KW cancer; cell proliferative disorder; autoimmune disorder;
KW inflammatory disorder; infection; hormonal disorder; metabolic disorder;
KW neurological disorder; gastrointestinal disorder; transport disorder;
KW connective tissue disorder; drug screening; proteome analysis;
KW gene therapy; antisense therapy; genotyping; transgenic animal; knock in;
KW disease model; toxicological testing; transcript imaging;
KW nucleic acid synthesis; nucleic acid modification.
XX
OS Homo sapiens.
XX
XX WO200297031-A2.
XX
XX 05-DEC-2002.
XX
XX 27-MAR-2002; 2002WO-US010056.
XX
XX 28-MAR-2001; 2001US-0279619P.
XX 29-MAR-2001; 2001US-0280067P.
XX 29-MAR-2001; 2001US-0280068P.
XX 16-MAY-2001; 2001US-0291280P.
XX 17-MAY-2001; 2001US-0291829P.
XX 17-MAY-2001; 2001US-0291849P.
XX 19-JUN-2001; 2001US-0299428P.
XX 20-JUN-2001; 2001US-0299776P.
XX 20-JUN-2001; 2001US-0300001P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshey SR;
PI Daughtery SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH;
PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
XX
XX WPI; 2003-129518/12.
XX
XX N-PSDB; ACC46487.
XX
XX Novel human diagnostic and therapeutic polypeptide useful for identifying
PT test compound which specifically binds to a polypeptide encoded by human
PT diagnostic and therapeutic polynucleotide, and to induce antibodies.
XX
XX Claim 27; SEQ ID NO 1084; 591pp; English.
XX
XX The invention relates to novel human diagnostic and therapeutic
CC polynucleotides designated dithp (ACC46080-ACC46749) and to their encoded
CC proteins (DITHP; ABR41136-ABR41812). The invention also relates to
CC polynucleotide sequences at least 90% identical to the dithp cDNA
CC sequences of the invention; recombinant vectors, host cells and
CC transgenic organisms comprising a dithp nucleic acid sequence; the
CC recombinant production of DITHP proteins; antibodies specific for DITHP
CC proteins; microarrays comprising dithp nucleic acid sequences; methods of
CC detecting dithp nucleotide and protein sequences; methods of screening
CC for compounds which specifically bind a DITHP protein; and methods of
CC assessing the toxicity of test compounds using a dithp hybridisation
CC probe. Dithp nucleic acid sequences and DITHP proteins may be used in the
CC diagnosis of a wide variety of conditions including cancer and other cell
CC proliferative disorders; autoimmune or inflammatory disorders; bacterial,
CC viral, fungal or parasitic infections; hormonal disorders; metabolic
CC disorders; neurological disorders; gastrointestinal disorders; transport
CC disorders; and connective tissue disorders. They may also be used to
CC screen for modulators of protein activity or gene expression. DITHP
CC proteins can additionally be used in analysis of the proteome of a tissue
CC or cell type and to induce antibodies. The dithp nucleic acids are
CC additionally useful in somatic or germline gene therapy of the disorders
CC mentioned above, as a source of antisense sequences, as a source of
CC probes and primers, in genotyping and identification of individuals, in
CC the generation of transgenic animal models of human disease or knock in
CC humanised animals, in toxicological testing, and in transcript imaging.
XX The present sequence represents a DITHP protein which is involved in
CC nucleic acid synthesis and/or modification. Note: The sequence data for
CC this patent did not form part of the printed specification, but was

CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 117 AA;
XX
XX Query Match 100.0%; Score 32; DB 6; Length 117;
XX Best Local Similarity 100.0%; Pred. No. 2.5e+02;
XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 CRGDD 5
XX |||||
XX 113 CRGDD 117
XX
XX RESULT 10
XX AAU44177
XX ID AAU44177 standard; protein; 120 AA.
XX
XX AC AAU44177;
XX
XX DT 27-FEB-2002 (first entry)
XX
XX DE Propionibacterium acnes immunogenic protein #5073.
XX
XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
XX Propionibacterium acnes.
XX
XX WO200181581-A2.
XX
XX 01-NOV-2001.
XX
XX 20-APR-2001; 2001WO-US012865.
XX
XX 21-APR-2000; 2000US-0199047P.
XX 02-JUN-2000; 2000US-0208841P.
XX 07-JUL-2000; 2000US-0216747P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
XX WPI; 2001-616774/71.
XX N-PSDB; AAS59522.
XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
XX Example 1; SEQ ID NO 5372; 1069pp; English.
XX
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was

CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 120 AA;

Query Match 100.0%; Score 32; DB 4; Length 120;

Best Local Similarity 100.0%; Pred. No. 2.5e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0;

QY 1 CRGDD 5

|||||

Db 105 CRGDD 109

RESULT 11

ABM40696

ID ABM40696 standard; protein; 120 AA.

XX
AC ABM40696;

XX
DT 20-OCT-2003 (first entry)

XX Propionibacterium acnes predicted ORF-encoded polypeptide #5372.

XX Acne vulgaris; antiseborrhoeic; dermatological; antibacterial;

XX immunostimulant; immune response; vaccine.

XX Propionibacterium acnes.

XX
WO2003033515-A1.

XX
24-APR-2003.

XX
11-OCT-2002; 2002WO-US032727.

XX
15-OCT-2001; 2001US-00978825.

XX
(CORI-) CORIXA CORP.

XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;

PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;

PI Barth B, Vallieue-Douglas J;

XX
WPI; 2003-381789/36.

DR N-PSDB; ACF64451.

XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.

XX
Example 1; SEQ ID NO 5372; 1481pp; English.

XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
XX encoding a Propionibacterium acnes protein. The invention also relates to
XX polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
XX immunogenic fragments of P. acnes polypeptides. The invention
XX additionally encompasses expression vectors and host cells comprising a
XX polynucleotide of the invention; antibodies against polypeptides of the
XX polynucleotide of the invention; fusion proteins comprising a polypeptide of the invention; a
XX method for stimulating an immune response specific for a P. acnes
XX polypeptide and an isolated T cell population comprising T cells prepared
XX via this method; a vaccine composition (comprising P. acnes polypeptides,
XX polynucleotides, antibodies, fusion proteins, T cell populations, or
XX antigen-presenting cells that express the polypeptide); a method and kit
XX for detecting or determining the presence or absence of P. acnes in a
XX patient; and a method for inhibiting the development of P. acnes in a
XX patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
XX proteins, T cell populations or antigen-presenting cells that express the
XX polypeptides are useful for diagnosing, preventing or treating acne
XX vulgaris, or for stimulating an immune response specific for a P. acnes
XX protein. The polynucleotides can also be used as probes or primers for
XX nucleic acid hybridisation. The vaccine composition is useful for the
XX stimulation of an immune response against P. acnes, or for treating acne,

CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 120 AA;

Query Match 100.0%; Score 32; DB 6; Length 120;

Best Local Similarity 100.0%; Pred. No. 2.5e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5

|||||

Db 105 CRGDD 109

RESULT 12

ADJ48934

ID ADJ48934 standard; protein; 152 AA.

XX
AC ADJ48934;

XX
DT 06-MAY-2004 (first entry)

XX Oil-associated gene related protein #434.

XX oil-associated gene; transgenic; enhanced seed oil; vegetable oil.

XX Unidentified.

XX
US2004025202-A1.

XX
05-FEB-2004.

XX
14-MAR-2003; 2003US-00389566.

XX
15-MAR-2002; 2002US-0365301P.

PR 26-JUN-2002; 2002US-0391786P.

PR 26-JUN-2002; 2002US-0392018P.

XX (LAUR/) LAURIE C C.

PA (RAVA/) RAVANELLO M.

PA (SAVA/) SAVAGE T.

PA (LEDE/) LEDEAUX J R.

PA (ROGE/) ROGERS J A.

XX Laurie CC, Ravanello M, Savage T, Ledeaux JR, Rogers JA;

XX
WPI; 2004-142683/14.

XX Novel recombinant DNA construct comprising a promoter functional in
PT plants operably linked to an oil-associated gene for producing transgenic
PT plant seed.

XX
Example 3; SEQ ID NO 938; 22pp; English.

XX The invention relates to a recombinant DNA construct comprising a
XX promoter functional in plants operably linked to an oil-associated gene.
XX The construct is useful for transgenic plant seed which has in its genome
XX the construct, that is functional in the plant to transcribe the oil-
XX associated gene. The transgenic plant seed grows into a plant having
XX enhanced seed oil as compared to wild type. The construct is useful for
XX producing hybrid maize seed. The transgenic plant seed is useful for
XX producing vegetable oil. The present sequence represents the amino acid
XX sequence of an oil-associated gene related protein.

XX
SQ Sequence 152 AA;

Query Match 100.0%; Score 32; DB 8; Length 152;

Best Local Similarity 100.0%; Pred. No. 3.2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

PR 01-MAY-2002; 2002US-00136826.
XX (CURA-) CURAGEN CORP.
XX
XX Miller CE, Kekuda R, Malyankar UM, Li L, Pena CBA, Spytek KA;
PI Gorman L, Guo X, Fernandes ER, Smithson G, Stone DJ, Zerhusen BD;
PI Patturajan M, Anderson DW, Mezes PS, Peyman JA, Macdougall JR;
PI Padigar M, Rastelli L, Shenoy SG, Gerlach VL, Shinkets RA, Zhong M;
PI Edinger SR, Ellerman K;
XX
DR WPI; 2003-239445/23.
DR N-PSDB; ADC39078.
XX
XX New NOVX polypeptides and polynucleotides, useful in gene therapy,
PT particularly for treating or preventing a syndrome associated with a
PT human disease e.g. diabetes, obesity, cancer, Alzheimer's disease,
PT hypertension or hemophilia.
XX
PS Claim 1; SEQ ID NO 18; 748pp; English.
XX
CC The invention relates to new isolated NOVX polypeptides, the genes
CC encoding them or sequences having at least 95% identity to the amino acid
CC or nucleotide sequences. The NOVX polypeptide is useful as a therapeutic,
CC particularly in the manufacture of a medicament for treating a syndrome
CC associated with a human disease, which includes a pathology associated
CC with NOVX polypeptide. The NOVX polypeptide is particularly useful for
CC treating, preventing or alleviating pathology associated with NOVX
CC polypeptide in a mammal, e.g. a human. The NOVX nucleic acid and
CC polypeptide are especially useful for treating or preventing e.g.
CC diabetes, obesity, cancers (e.g. lymphoma, uterus cancer or prostate
CC cancer), dyslipidemias, anorexia, wasting disorders, Alzheimer's disease,
CC Parkinson's disorder, cachexia, cardiomyopathy, AIDS, asthma, Crohn's
CC disease, multiple sclerosis, hypertension, atherosclerosis, hemophilia,
CC graft-versus-host disease or Albrit hereditary osteodystrophy. The DNA
CC encoding the protein is useful in gene therapy for treating the above
CC conditions. These are also useful in developing powerful assay system for
CC functional analysis of various human disorders, as well as in diagnostic
CC applications. This sequence represents one of the NOVX proteins of the
CC invention.
XX
SQ Sequence 156 AA;
Query Match 100.0%; Score 32; DB 7; Length 156;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CRGDD 5
Db 62 CRGDD 66
|||||
ADSD26705
ID ADS26705 standard; protein; 301 AA.
XX
AC ADS26705;
XX
XX 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #15738.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
OS Bacteria.
XX
XX US2003233675-A1.
XX

PD 18-DEC-2003.
XX
XX 20-FEB-2003; 2003US-00369493.
XX
XX 21-FEB-2002; 2002US-0360039P.
XX
XX (CAOV/) CAO Y.
XX (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
PI WPI; 2004-061375/06.
DR
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 15738; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition. Improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 301 AA;
Query Match 100.0%; Score 32; DB 8; Length 301;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CRGDD 5
Db 200 CRGDD 204
|||||
ADSD27091
ID ADS27091 standard; protein; 301 AA.
XX
AC ADS27091;
XX
XX 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #16124.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW bacterial polypeptide.

KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.

XX Bacteria.

XX US2003233675-A1.

XX 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.

XX (CAOY/) CAO Y.

XX (HINK/) HINKLE G J.

XX (SLAT/) SLATER S C.

XX (CHEN/) CHEN X.

XX (GOLD/) GOLDMAN B S.

XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI; 2004-061375/06.

XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.

XX Claim 1; SEQ ID NO 16124; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 301 AA;

Query Match 100.0%; Score 32; DB 8; Length 301;

Best Local Similarity 100.0%; Pred. No. 5.9e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5

|||||

200 CRGDD 204

RESULT 17

ADS26338

ID ADS26338 standard; protein; 302 AA.

XX ADS26338;

XX 02-DEC-2004 (first entry)

XX

DE Bacterial polypeptide #15371.

XX Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.

XX Bacteria.

XX US2003233675-A1.

XX 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.

XX (CAOY/) CAO Y.

XX (HINK/) HINKLE G J.

XX (SLAT/) SLATER S C.

XX (CHEN/) CHEN X.

XX (GOLD/) GOLDMAN B S.

XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI; 2004-061375/06.

XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.

XX Claim 1; SEQ ID NO 15371; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 302 AA;

Query Match 100.0%; Score 32; DB 8; Length 302;

Best Local Similarity 100.0%; Pred. No. 5.9e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5

|||||

203 CRGDD 207

RESULT 18

AAW79286
 ID AAW79286 standard; protein; 311 AA.
 XX AAW79286;
 AC
 XX 15-FEB-1999 (first entry)
 XX Human phosphatidic acid phosphatase beta.
 DE
 XX Phosphatidic acid phosphatase beta; PAP-beta; human; dephosphorylation;
 KW tumour suppressor; cancer; gene therapy.
 KW
 XX Homo sapiens.
 OS
 XX WO9846730-A1.
 PN
 XX 22-OCT-1998.
 PD
 XX 16-APR-1998; 98WO-US007928.
 XX
 PF 17-APR-1997; 97US-00842827.
 PR
 XX (CELL-) CELL THERAPEUTICS INC.
 PA
 XX Leung DW, Tompkins CK;
 PI
 XX WPI; 1998-594568/50.
 DR
 XX N-PSDB; AAV69088.
 XX
 XX New nucleic acid encoding human phosphatidic acid phosphatases - used to
 PT regulate levels of lipid cellular mediators and in gene therapy of e.g.
 PT cancer.
 PT
 XX Claim 8; Fig 3A-B; 62pp; English.
 PS
 XX This is the amino acid sequence of human mature phosphatidic acid
 CC phosphatase-beta (PAP-beta), an enzyme that catalyses the conversion of
 CC phosphatidic acid into diacylglycerol. 3 Variants of human PAP, i.e. PAP-
 CC alpha 1 and 2 (see AAW79284-85), PAP-beta (see AAW79286) and PAP-gamma
 CC (see AAW79287) have been identified. The invention provides PAP
 CC polynucleotides (see AAV69086-89) and polypeptides, a method of preparing
 CC PAP in a transformed host cell, and a method of using PAP to
 CC dephosphorylate a substrate, especially lysophosphatidic acid, ceramide 1
 CC -phosphate or sphingosine 1-phosphate, particularly for production of
 CC diacylglycerol, but also monoacylglycerol, ceramide and sphingosine. PAP
 CC is able to control the balance of lipid mediators of cellular activation
 CC and signal transduction. Sequences that encode PAP are potentially
 CC useful, in gene therapy, for treatment of cancer (PAP may be tumour
 CC suppressors, PAP-alpha is expressed at lower levels in cancer cells than
 CC in normal cells of same tissue type), inflammatory disease and diabetes-
 CC associated obesity
 XX Sequence 311 AA;
 SQ
 Query Match 100.0%; Score 32; DB 2; Length 311;
 Best Local Similarity 100.0%; Pred. No. 6.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CRGDD 5
 DB 181 CRGDD 185
 RESULT 19
 ADE60703
 ID ADE60703 standard; protein; 311 AA.
 XX
 AC ADE60703;
 XX
 XX 29-JAN-2004 (first entry)
 DT
 XX Human Protein NP_003704, SEQ ID NO 6615.
 DE
 XX

KW Human; pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; chronic constriction injury; CCI;
 KW spared nerve injury; SNI; Chung.
 XX Homo sapiens.
 OS
 XX WO2003016475-A2.
 PN
 XX 27-FEB-2003.
 PD
 XX 14-AUG-2002; 2002WO-US025765.
 PF
 XX 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX
 XX (GEHO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 XX
 XX Woolf C, D'urso D, Befort K, Costigan M;
 PI
 XX WPI; 2003-268312/26.
 DR GENBANK; NP_003704.
 DR
 XX New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 PT
 XX Claim 1; Page; 1017pp; English.
 PS
 XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a human protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 311 AA;
 SQ
 Query Match 100.0%; Score 32; DB 7; Length 311;
 Best Local Similarity 100.0%; Pred. No. 6.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CRGDD 5
 DB 181 CRGDD 185
 RESULT 20
 ADM72103
 ID ADM72103 standard; protein; 311 AA.
 XX
 XX ADM72103;
 AC

XX DT 03-JUN-2004 (first entry)
 XX PA Human tau-related polypeptide PPAP2B.
 XX DE
 XX KW Tau; PPAP2A; PPAP2B; IHPK1; IHPK3; IHPK2; FLJ20530; DJ434014.5; EZF1;
 KW Loc127424; Loc113179; KIAA0826; neuroprotective; neurotropic;
 KW antiparkinsonian; tau-protein kinase; human.
 XX OS Homo sapiens.
 XX PN WO2004022708-A2.
 XX PD 18-MAR-2004.
 XX PF 04-SEP-2003; 2003WO-US027590.
 XX PR 09-SEP-2002; 2002US-040877P.
 XX PA (BGHM) BRIGHAM & WOMENS HOSPITAL INC.
 XX PI Feany MB, Shulman JM;
 XX DR WPI; 2004-248456/23.
 XX KW New substantially pure Tau-related polypeptides and polynucleotides,
 PT useful for diagnosing and/or treating neurological conditions with
 PT aberrant expression of the Tau-related polypeptide, such as Alzheimer's
 PT disease.
 XX PS Claim 1; SEQ ID NO 2; 62pp; English.
 XX CC The invention relates to tau-related polypeptide consisting essentially
 CC of an amino acid sequence selected from PPAP2A, PPAP2B, IHPK1, IHPK3,
 CC IHPK2, FLJ20530, DJ434014.5, EZF1, Loc127424, Loc113179 and KIAA0826
 CC (ADM72102-ADM72112 respectively). The methods and compositions of the
 CC present invention are useful for the diagnosis and/or treatment of
 CC neurological diseases or conditions associated with aberrant expression
 CC or activity of the Tau-related polypeptide, such as Alzheimer's disease
 CC and Parkinson's disease. The present sequence represents a human tau-
 CC related polypeptide homologue PPAP2B.
 XX SQ Sequence 311 AA;
 Query Match 100.0%; Score 32; DB 8; Length 311;
 Best Local Similarity 100.0%; Pred. No. 6.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CRGDD 5
 DB 181 CRGDD 185
 RESULT 21
 ID ADN04415 standard; protein; 311 AA.
 XX AC ADN04415;
 XX DT 01-JUL-2004 (first entry)
 XX DE Antipsoriatic protein sequence #401.
 XX KW antipsoriatic; gene therapy; psoriasis; diagnosis.
 XX OS Homo sapiens.
 XX PN WO2004028479-A2.
 XX PD 08-APR-2004.
 XX PF 25-SEP-2003; 2003WO-US030907.
 XX CC

PR 25-SEP-2002; 2002US-0414006P.
 XX (GETH) GENENTECH INC.
 XX PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
 PI Wu TD;
 XX DR WPI; 2004-305105/28.
 DR N-PSDB; ADN04414.
 XX PT New PRO nucleic acid or polypeptide, useful for preparing a
 PT pharmaceutical composition for diagnosing or treating psoriasis in a
 PT mammal.
 XX PS Claim 9; SEQ ID NO 809; 3069pp; English.
 XX CC The invention relates to novel polynucleotide and polypeptides for
 CC treating psoriasis or a sequence having at least 80% identity to the
 CC above sequences. The nucleic acid is useful for preparing a composition
 CC for diagnosing or treating psoriasis in a mammal. This sequence
 CC corresponds to one of the polypeptides of the invention.
 XX SQ Sequence 311 AA;
 Query Match 100.0%; Score 32; DB 8; Length 311;
 Best Local Similarity 100.0%; Pred. No. 6.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CRGDD 5
 DB 181 CRGDD 185
 RESULT 22
 ADP24274
 ID ADP24274 standard; protein; 311 AA.
 XX AC ADP24274;
 DT 18-NOV-2004 (first entry)
 XX DE PRO polypeptide SEQ ID NO:1452.
 XX KW PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;
 KW osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;
 KW antiasthmatic; hepatotropic; respiratory; gene therapy; immune system.
 XX OS Unidentified.
 XX PN WO2004041170-A2.
 XX PD 21-MAY-2004.
 XX PF 30-OCT-2003; 2003WO-US034312.
 XX PR 01-NOV-2002; 2002US-0423394P.
 XX PA (GETH) GENENTECH INC.
 XX PI Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;
 PI Wu TD;
 XX DR WPI; 2004-419628/39.
 DR N-PSDB; ADP24273.
 XX PT New PRO polypeptides and polynucleotides, useful for treating e.g.
 PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated
 PT renal disease, or demyelinating diseases of the central or peripheral
 PT nervous system.
 XX PS Claim 7; SEQ ID NO 1452; 2940pp; English.
 XX CC The invention relates to a novel isolated nucleic acid and the PRO

CC polypeptide encoded by it. A protein of the invention has
 CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
 CC osteopathic, antidiabetic, dermatological, antiporiatic, antiallergic,
 CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide
 CC of the invention may have a use in gene therapy. The PRO polypeptide, its
 CC agonist, antagonist, or antibody that specifically binds to the
 CC polypeptide is useful for treating an immune related disorder such as
 CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
 CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
 CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
 CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
 CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
 CC disease, a demyelinating disease of the central or peripheral nervous
 CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
 CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
 CC disease, infectious or autoimmune chronic active hepatitis, primary
 CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
 CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
 CC disease, an autoimmune or immune-mediated skin disease, a bullous skin
 CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
 CC disease, asthma, allergic rhinitis, atopic dermatitis, food
 CC hypersensitivity, urticaria, an immunologic disease of the lung,
 CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
 CC pneumonitis, a transplantation associated disease, graft rejection or
 CC graft-versus-host disease. The present sequence represents a PRO protein
 CC of the invention.

XX
 SQ Sequence 311 AA;

Query Match 100.0%; Score 32; DB 8; Length 311;
 Best Local Similarity 100.0%; Pred. No. 6.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
 |||||
 DB 181 CRGDD 185

RESULT 23

ABO62272
 ID ABO62272 standard; protein; 376 AA.

XX
 AC ABO62272;

XX
 DT 29-JUL-2004 (first entry)

XX Klebsiella pneumoniae polypeptide seqid 8789.

XX
 KW Recombinant expression vector; transcription regulatory element;
 XX Klebsiella pneumoniae protein; antibacterial; Vaccine.

XX Klebsiella pneumoniae.

PN US6610836-B1.

XX
 PD 26-AUG-2003.

XX
 PF 27-JAN-2000; 2000US-00489039.

XX
 PR 29-JAN-1999; 99US-0117747P.

XX
 PA (GENO-) GENOME THERAPEUTICS CORP.

XX Breton GL, Osborne M;

XX
 DR WPI; 2003-895346/82.

XX
 DR N-PSDB; ACH95823.

XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
 PT preparing a vaccine composition against Klebsiella pneumoniae.

XX
 PS Disclosure; SEQ ID NO 8789; 932pp; English.

XX

CC The invention describes a new isolated nucleic acid encoding a Klebsiella
 CC pneumoniae polypeptide. Also described are: a recombinant expression
 CC vector comprising the nucleic acid, operably linked to a transcription
 CC regulatory element; and a cell comprising the recombinant expression
 CC vector. The nucleic acid is useful for preparing a vaccine composition
 CC against Klebsiella pneumoniae. This is the amino acid sequence of a
 CC Klebsiella pneumoniae polypeptide of the invention

XX Sequence 376 AA;

Query Match 100.0%; Score 32; DB 7; Length 376;
 Best Local Similarity 100.0%; Pred. No. 7.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
 |||||
 DB 158 CRGDD 162

RESULT 24

ABU31582

ID ABU31582 standard; protein; 400 AA.

XX
 AC ABU31582;

XX
 DT 19-JUN-2003 (first entry)

XX Protein encoded by Prokaryotic essential gene #17109.

XX
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Klebsiella pneumoniae.

PN WO200277183-A2.

XX
 PD 03-OCT-2002.

XX
 PF 21-MAR-2002; 2002WO-US009107.

XX
 PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

DR N-PSDB; ACA35452.

XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 59506; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological

CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 400 AA;

Query Match 100.0%; Score 32; DB 6; Length 400;
 Best Local Similarity 100.0%; Pred. No. 7.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
 |||||
 Db 41 CRGDD 45

RESULT 25
 ABU45262
 ID ABU45262 standard; protein; 400 AA.
 AC ABU45262;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #30789.
 XX
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS Salmonella paratyphi.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

DR N-PSDB; ACA49132.

XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 73186; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid

CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 400 AA;

Query Match 100.0%; Score 32; DB 6; Length 400;
 Best Local Similarity 100.0%; Pred. No. 7.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
 |||||
 Db 41 CRGDD 45

RESULT 26

ABU48051
 ID ABU48051 standard; protein; 400 AA.

XX AC ABU48051;

XX 19-JUN-2003 (first entry)

XX Protein encoded by Prokaryotic essential gene #33578.

DE Antisense; prokaryotic essential gene; cell proliferation; drug design.

OS Salmonella typhi.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

DR N-PSDB; ACA51921.

XX

PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

XX
XX
XX Claim 25; SEQ ID NO 75975; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX
XX Sequence 400 AA;

Query Match 100.0%; Score 32; DB 6; Length 400;
Best Local Similarity 100.0%; Pred. No. 7.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
|||||
Db 41 CRGDD 45

RESULT 27
ABU15369
ID ABU15369 standard; protein; 400 AA.
XX
XX ABU15369;
XX
XX
DT 19-JUN-2003 (first entry)
XX
XX Protein encoded by Prokaryotic essential gene #896.
DE
DE Antisense; prokaryotic essential gene; cell proliferation; drug design.
KW
XX Escherichia coli.
OS
XX WO200277183-A2.
XX
XX
XX 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US0009107.
XX
XX 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.
XX
XX (SLIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI; 2003-029926/02.
DR N-PSDB; ACA19239.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

XX
XX Claim 25; SEQ ID NO 43293; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX
XX Sequence 400 AA;

Query Match 100.0%; Score 32; DB 6; Length 400;
Best Local Similarity 100.0%; Pred. No. 7.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
|||||
Db 41 CRGDD 45

RESULT 28
ABO65513
ID ABO65513 standard; protein; 424 AA.
XX
XX ABO65513;
XX
XX 29-JUL-2004 (first entry)
XX
XX Klebsiella pneumoniae polypeptide seqid 12030.
DE
XX Recombinant expression vector; transcription regulatory element;
KW Klebsiella pneumoniae protein; antibacterial; Vaccine.
XX
XX Klebsiella pneumoniae.

XX US6610836-B1.
 XX
 XX 26-AUG-2003.
 XX
 XX 27-JAN-2000; 2000US-00489039.
 XX
 XX 29-JAN-1999; 99US-0117747P.
 XX
 XX (GENO-) GENOME THERAPEUTICS CORP.
 XX
 XX Breton GL, Osborne M;
 XX
 XX WPI; 2003-895346/82.
 XX
 XX N-PSDB; ACH99064.
 XX
 XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
 PT preparing a vaccine composition against Klebsiella pneumoniae.
 XX
 XX Disclosure; SEQ ID NO 12030; 932pp; English.
 XX
 XX The invention describes a new isolated nucleic acid encoding a Klebsiella
 CC pneumoniae polypeptide. Also described are: a recombinant expression
 CC vector comprising the nucleic acid, operably linked to a transcription
 CC regulatory element; and a cell comprising the recombinant expression
 CC vector. The nucleic acid is useful for preparing a vaccine composition
 CC against Klebsiella pneumoniae. This is the amino acid sequence of a
 CC Klebsiella pneumoniae polypeptide of the invention
 CC
 XX Sequence 424 AA;
 SQ

Query Match 100.0%; Score 32; DB 7; Length 424;
 Best Local Similarity 100.0%; Pred. No. 8.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CRGDD 5
 Db 65 CRGDD 69
 |||||

RESULT 29
 ABO66647
 ID ABO66647 standard; protein; 488 AA.
 XX
 XX ABO66647;
 XX
 XX 29-JUL-2004 (first entry)
 DT
 XX
 XX Klebsiella pneumoniae polypeptide seqid 13164.
 DE
 XX
 XX Recombinant expression vector; transcription regulatory element;
 KW Klebsiella pneumoniae protein; antibacterial; Vaccine.
 KW
 XX Klebsiella pneumoniae.
 OS
 XX
 XX US6610836-B1.
 XX
 XX 26-AUG-2003.
 XX
 XX 27-JAN-2000; 2000US-00489039.
 XX
 XX 29-JAN-1999; 99US-0117747P.
 XX
 XX (GENO-) GENOME THERAPEUTICS CORP.
 XX
 XX Breton GL, Osborne M;
 XX
 XX WPI; 2003-895346/82.
 XX
 XX N-PSDB; ABO00218.
 XX
 XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
 PT preparing a vaccine composition against Klebsiella pneumoniae.
 XX

PS Disclosure; SEQ ID NO 13164; 932pp; English.
 XX
 XX The invention describes a new isolated nucleic acid encoding a Klebsiella
 CC pneumoniae polypeptide. Also described are: a recombinant expression
 CC vector comprising the nucleic acid, operably linked to a transcription
 CC regulatory element; and a cell comprising the recombinant expression
 CC vector. The nucleic acid is useful for preparing a vaccine composition
 CC against Klebsiella pneumoniae. This is the amino acid sequence of a
 CC Klebsiella pneumoniae polypeptide of the invention
 XX
 XX Sequence 488 AA;
 SQ

Query Match 100.0%; Score 32; DB 7; Length 488;
 Best Local Similarity 100.0%; Pred. No. 9.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CRGDD 5
 Db 103 CRGDD 107
 |||||

RESULT 30
 ADN19416
 ID ADN19416 standard; protein; 1116 AA.
 XX
 XX ADN19416;
 AC
 XX 02-DEC-2004 (first entry)
 DT
 XX
 XX Bacterial polypeptide #2069.
 DE
 XX
 XX Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.
 KW
 XX Bacteria.
 OS
 XX US2003233675-A1.
 XX
 XX 18-DEC-2003.
 XX
 XX 20-FEB-2003; 2003US-00369493.
 XX
 XX 21-FEB-2002; 2002US-0360039P.
 XX
 XX (CAOY) CAO Y.
 PA (HINK) HINKLE G J.
 PA (SLAT) SLATER S C.
 PA (CHEN) CHEN X.
 PA (GOLD) GOLDMAN B S.
 XX
 XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 PI
 XX WPI; 2004-061375/06.
 XX
 XX New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX
 XX Claim 1; SEQ ID NO 2069; 122pp; English.
 PS
 XX The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the

CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 SQ Sequence 1116 AA;

Query Match 100.0%; Score 32; DB 8; Length 1116;
 Best Local Similarity 100.0%; Pred. No. 26+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
 |||||
 Db 677 CRGDD 681

RESULT 31
 ABB62913
 ID ABB62913 standard; protein; 1837 AA.
 XX
 AC ABB62913;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 15531.

XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 KW Drosophila melanogaster.

OS WO200171042-A2.
 PN 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.
 XX 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.
 XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI; 2001-656860/75.

XX N-PSDB; ABL07016.
 DR

XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.

PS Disclosure; SEQ ID NO 15531; 21pp + Sequence Listing; English.
 XX

CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS57737-
 CC ABB57072). The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1837 AA;

Query Match 100.0%; Score 32; DB 4; Length 1837;
 Best Local Similarity 100.0%; Pred. No. 3.1e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
 |||||
 Db 1413 CRGDD 1417

RESULT 32
 AAR96407
 ID AAR96407 standard; peptide; 7 AA.

XX AAR96407;

XX 21-OCT-2004 (revised)
 DT 05-JUL-1996 (first entry)

XX RGD cyclic peptide, TL#114, binds alpha-v,beta3 integrin receptor.

XX RGD-containing peptide; alpha-v, beta-3 integrin receptor; osteoclast;
 KW matrix; bone; inhibition; bone resorption; promote; endothelial cell;
 KW smooth muscle cell; restenosis; angiogenesis; cyclic.

XX Synthetic.

XX Key Location/Qualifiers
 FH Disulfide-bond 1..7

FT Modified-site 1 /note= "Acyl-Cys"

FT Modified-site 7 /note= "Amidated C-terminal"

XX WO9528426-A2.

XX 26-OCT-1995.

XX 12-APR-1995; 95WO-US004741.

XX 13-APR-1994; 94US-00227316.

PR 08-SEP-1994; 94US-00303052.

XX (LJOL-) LA JOLLA CANCER RES FOUND.

XX Cheng S, Ingram R, Mullen D, Tschopp J;

XX WPI; 1995-373767/48.

XX Altering alpha-v, beta-3 integrin receptor-mediated binding of cell to
 PT matrix - using conformationally restrained peptide of RGD type, e.g. for
 PT treating inappropriate angiogenesis or for inhibiting bone resorption.

XX Claim 85; Page 80; 99pp; English.

XX The sequences given in AAR96301-417 are non-naturally occurring RGD-
 CC containing peptides which alter the alpha-v, beta-3 integrin receptor
 CC binding of a cell to a matrix, such as the binding of an osteoclast to a
 CC matrix such as bone. These peptides inhibit bone resorption and can
 CC inhibit or promote alpha-v, beta-3-mediated cell attachment depending on
 CC whether they are present in the cell in a soluble form or are bound to a
 CC solid substrate. These peptides can be used in the amelioration of the
 CC severity of a pathology involving alpha-v, beta-3 receptor-mediated
 CC binding of a cell, such as an osteoclast, endothelial cell or smooth
 CC muscle cell to a matrix. They are used for treating conditions associated
 CC with restenosis or inappropriate or insufficient angiogenesis, or for
 CC inhibiting osteoclast binding to the matrix

CC Revised record issued on 21-OCT-2004 : Correction to Feature Table Key

```
XX SQ Sequence 7 AA;
Query Match 90.6%; Score 29; DB 2; Length 7;
Best Local Similarity 80.0%; Pred. No. 1.8e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 CRGDD 5
Db 1 CKGDD 5

RESULT 33
AAW66822
ID AAW66822 standard; peptide; 7 AA.
XX AC AAW66822;
XX DT 25-MAR-2003 (revised)
DT 10-DEC-1998 (first entry)
DE Peptide useful for altering bone resorption.
XX bone resorption; pharmacore; angiogenesis; restenosis; integrin receptor;
KW alpha v beta 3 integrin receptor; osteoclast; cyclic.
XX Synthetic.
XX FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N-terminal acetyl"
FT Modified-site 7
FT /note= "C-terminal amide"
XX US5807819-A.
XX PD 15-SEP-1998.
XX PF 12-APR-1995; 95US-00421698.
XX PR 15-APR-1994; 94US-00227316.
PR 08-SEP-1994; 94US-00303052.
XX (LJOL-) LA JOLLA CANCER RES CENT.
XX TSchoopp JF, Mullen D, Ingram R, Cheng S;
XX WPI; 1998-555601/47.
XX Use of peptide derivatives which can alter integrin receptor binding -
PT for altering bone resorption, treating angiogenesis or restenosis and
PT altering integrin receptor mediated interactions.
XX
XX Example 2; Fig 2D; 87pp; English.
XX A new method is claimed for altering bone resorption. It comprises
CC administration of a peptide of formula: X1X2X3X4GX5X6X7X8; where X1 =
CC R1R2N or 0-10 amino acids (optionally protected by acetylation at the N-
CC terminus); X2 = absent or 1 amino acid; X3 = absent or 1 or 2 amino acids
CC ; X4 = N-Me-Arg; X5 = residue which provides an ionic interaction with an
CC integrin receptor, or is Mea, Pea or Tfsa; X6 = residue which has an
CC aliphatic side chain; a non-natural amino acid that is hydrophobic; or
CC Thr; X7 = a residue capable of forming a bond (i) with a bridging amino
CC acid of X2, (ii) with X3 when X2 is absent, or (iii) with X4 when X2 and
CC X3 are absent, to conformationally restrain the peptide; X8 = NR3R4; OR5;
CC or 0-10 amino acids, optionally protected as an amide at the C- terminus;
CC R1, R3-R5 = H or alkyl; R2 = H, alkyl, alkyl-CO or phenyl-CO. The
CC peptides are useful for inhibiting bone resorption, angiogenesis or
CC restenosis, and for altering integrin receptor-mediated interactions,
CC especially alpha v beta 3 integrin receptor-mediated binding of cells to
CC a matrix. They may be used for reducing or inhibiting osteoclast binding
CC to a matrix. The present sequence represents an example of a circular
CC peptide disclosed in the specification. (Updated on 25-MAR-2003 to

CC correct PF field.)
XX SQ Sequence 7 AA;
Query Match 90.6%; Score 29; DB 2; Length 7;
Best Local Similarity 80.0%; Pred. No. 1.8e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 CRGDD 5
Db 1 CKGDD 5

RESULT 34
AAW97002
ID AAW97002 standard; peptide; 7 AA.
XX AC AAW97002;
XX DT 26-APR-1999 (first entry)
DE Peptide used to inhibit bone resorption, angiogenesis and restenosis.
XX RGD-peptide; bone resorption; angiogenesis; restenosis;
KW integrin receptor mediated interaction; binding; osteoclast; bone.
XX Synthetic.
XX FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "acetylated residue"
FT Modified-site 7
FT /note= "amidated residue"
XX US5849865-A.
XX PD 15-DEC-1998.
XX PF 12-APR-1995; 95US-00421695.
XX PR 13-APR-1994; 94US-00227316.
PR 08-SEP-1994; 94US-00303052.
XX (LJOL-) LA JOLLA CANCER RES FOUND.
XX TSchoopp JF, Mullen D, Ingram R, Cheng S;
XX WPI; 1999-179486/15.
XX RGD peptides - that modulate integrin-mediated cell-matrix binding.
XX Claim 6; Col 138; 90pp; English.
XX The present peptide represents a non-naturally occurring Arg-Gly-Asp
CC (RGD)-containing peptide. The peptide may be useful for reducing or
CC inhibiting bone resorption, angiogenesis or restenosis, and for altering
CC an integrin receptor mediated interaction. The peptide may also be used
CC to inhibit the binding of an osteoclast to bone
XX SQ Sequence 7 AA;
Query Match 90.6%; Score 29; DB 2; Length 7;
Best Local Similarity 80.0%; Pred. No. 1.8e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 CRGDD 5
Db 1 CKGDD 5

RESULT 35
AAW87384
ID AAW87384 standard; protein; 41 AA.
```

XX
AC AAM87384;
XX
XX 07-NOV-2001 (first entry)
XX
XX Human immune/haematopoietic antigen SEQ ID NO:14977.
DE
DE
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis.
XX
OS Homo sapiens.
XX
XX WO200157182-A2.
PN
XX
XX 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001354.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 13-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-02556719P.

CC complement, degenerate variant or fragment; (3) a genetically engineered
 CC host cell, transfected, transformed or infected with the vector of (2);
 CC (4) an antibody specific for the polypeptide of (1); (5) an immunogenic
 CC composition comprising the polypeptide, its complement, biological
 CC equivalent or fragment, or the polynucleotide that is comprised in the
 CC expression vector; (6) a pharmaceutical composition comprising the
 CC polypeptide of (1) and a carrier; (7) a protein chip comprising an array
 CC of the polypeptides of (1), their biological equivalent or fragment; (8)
 CC immunising against *Alloiooccus* otitis by administering to a host the
 CC otitis in the biological sample; (10) a kit comprising a container
 CC containing the novel polynucleotide, its degenerate variant or fragment,
 CC or the antibody of (4); and (11) producing a polypeptide by culturing the
 CC genetically engineered host cell under conditions suitable to produce the
 CC polypeptide from the culture. (1) can be used in gene therapy. The
 CC polynucleotides, polypeptides, antibodies and compositions of the present
 CC invention can be used for treating and diagnosing diseases, drug
 CC screening assays and monitoring of effects during drug clinical trials.
 CC The polynucleotides are useful for expressing and detecting *Alloiooccus*
 CC otitis. The present sequence represents an *Alloiooccus* otitis
 CC antigen protein from the present invention.

XX
 SQ Sequence 274 AA;

Query Match 90.6%; Score 29; DB 6; Length 274;
 Best Local Similarity 80.0%; Pred. No. 1.9e+03;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
 |:|:|
 Db 115 CKGDD 119

RESULT 38
 ADJ71969
 ID ADJ71969 standard; protein; 299 AA.
 XX
 AC ADJ71969;
 XX
 XX 06-MAY-2004 (first entry)
 XX Human PMMM protein amino acid sequence SeqID29.
 XX
 KW protein modification and maintenance molecule; PMMM; cytostatic;
 KW antiarteriosclerotic; anticonvulsant; nootropic; neuroprotective;
 KW cerebroprotective; vasotropic; anti-HIV; antiallergic; antimicrobial;
 KW antiinflammatory; endocrine-Gen; thyromimetic;
 KW cell proliferative disorder; cancer; atherosclerosis;
 KW neurological disorder; epilepsy; Huntington's disease; stroke;
 KW immune disorder; inflammatory disorder; AIDS; allergy;
 KW developmental disorder; Hypothyroidism; Cushing's syndrome; infection;
 KW human.

OS Homo sapiens.
 XX
 XX WO2004009797-A2.
 XX
 XX 29-JAN-2004.
 XX
 XX 23-JUL-2003; 2003WO-US023249.
 XX
 XX 23-JUL-2002; 2002US-0398143P.
 XX 09-AUG-2002; 2002US-0402458P.
 XX 12-AUG-2002; 2002US-0403289P.
 XX 27-AUG-2002; 2002US-0406472P.
 XX 06-SEP-2002; 2002US-0409354P.
 XX
 XX (INCY-) INCYTE CORP.

XX Elliott VS, Swarnakar A, Griffin JA, Lee EA, Sprague WW;
 PI Hafalia AJA, Lee SY, Kable AE, Ison CH, Khare R, Chawla NK;
 PI Marquis JP, Jiang X, Jackson AA, Becha SD, Emerling BM, Jin P;
 PI Wilson AD, Richardson TW, Yang J, Baughn MR, Gandhi AR, Nguyen DB;

PI Ramkumar J, Kallick DA, Kearney L, Lu DAM, Gietzen KJ, Tribouley CM;
 PI Lal PG, Blake JJ, Lu Y, Arvizu CS;
 XX
 DR WPI; 2004-123392/12.
 DR N-PSDB; ADJ72031.
 XX
 XX New human protein modification and maintenance molecules (PMMM), useful
 PT for diagnosing, treating and preventing diseases or conditions associated
 PT with the aberrant PMMM expression e.g. cancer, AIDS, epilepsy, or
 PT infections.

XX Claim 1; SEQ ID NO 29; 387pp; English.
 XX
 CC This invention relates to novel protein modification and maintenance
 CC molecules (PMMM) and polynucleotides which identify and encode PMMM. The
 CC invention may be useful for the development of compositions with a
 CC cytostatic, antiarteriosclerotic, anticonvulsant, nootropic,
 CC neuroprotective, cerebroprotective, vasotropic, anti-HIV, antiallergic,
 CC antimicrobial, antiinflammatory, endocrine-Gen or thyromimetic activity.
 CC The polypeptides and polynucleotides are useful in diagnosing, treating
 CC and preventing diseases or conditions associated with the decreased
 CC expression or overexpression of PMMM, such as cell proliferative (for
 CC example cancer, atherosclerosis), neurological (for example epilepsy,
 CC Huntington's disease, stroke), immune/inflammatory (for example AIDS,
 CC allergies) and developmental (for example Hypothyroidism, Cushing's
 CC syndrome) disorders, or infections. These are also useful in assessing
 CC the effects of exogenous compounds on the expression of nucleic acid and
 CC amino acid sequences of PMMM. The present sequence is that of a human
 CC PMMM protein of the invention.

XX Sequence 299 AA;

Query Match 90.6%; Score 29; DB 8; Length 299;
 Best Local Similarity 80.0%; Pred. No. 2.1e+03;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
 |:|:|
 Db 261 CKGDD 265

RESULT 39
 ADA50457
 ID ADA50457 standard; protein; 357 AA.
 XX
 AC ADA50457;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 XX Human protease SEQ ID NO:55.
 DE enzyme; human; protease.
 XX
 XX Homo sapiens.
 OS
 XX WO2003040393-A2.
 PN
 XX 15-MAY-2003.
 XX
 XX 04-NOV-2002; 2002WO-IB004615.
 XX
 XX 06-NOV-2001; 2001US-0332633P.
 XX
 XX (DECO-) DECODE GENETICS EHF.
 XX
 XX Martinez RAM, Sigurdson GT;
 PI
 XX WPI; 2003-441582/41.
 DR
 DR N-PSDB; ADA50504.

XX Novel isolated protease polypeptide and polynucleotide encoding the
 PT polypeptide useful for diagnosing and treating diseases or conditions
 PT associated with a protease.

XX Claim 1; Page 66; 160pp; English.

XX The invention relates to a novel isolated polypeptide comprising an amino

CC acid sequence that has greater than 95 % identity to any one of 47 150-

CC 350 residue protease polypeptide sequences, given in the specification.

CC The nucleic acids, probes, primers, polypeptides and antibodies of the

CC invention can be used in methods of diagnosis of a susceptibility to a

CC disease or condition associated with a protease. The present sequence

CC represents a protease of the invention.

XX Sequence 357 AA;

SQ

Query Match 90.6%; Score 29; DB 6; Length 357;

Best Local Similarity 80.0%; Pred. No. 2.5e+03;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5

Db 113 CRGDD 117

RESULT 40

ABR40851

ID ABR40851 standard; protein; 412 AA.

XX AC ABR40851;

XX DT 16-MAY-2003 (first entry)

XX DE Glycine max oil trait related protein sequence SEQ ID NO:481.

XX Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;

KW receptor-like protein kinase; mitogen activated protein kinase; oil;

KW LIP15-like transcription factor caleosin; ATP citrate lyase; SNF1;

KW CKC-like transcription factor; antisense inhibition; co-suppression;

transgenic plant.

XX Glycine max.

OS

XX WO2003002751-A2.

PN

XX 09-JAN-2003.

PD

XX 27-JUN-2002; 2002WO-US020152.

PF

XX 29-JUN-2001; 2001US-0301913P.

PR

XX (DUPO) DU PONT DE NEMOURS & CO E I.

PA (PION-) PIONEER HI-BRED INT INC.

PA Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harvell LT;

PI Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B;

PI Tarczyński MC;

PI WPI; 2003-201509/19.

DR N-PSDB; ACC00845.

DR

XX Novel nucleotide fragment encoding polypeptides having receptor-like

PT protein kinase activity, caleosin-like activity, useful for altering oil

PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.

XX Claim 12; Page 497-498; 542pp; English.

XX The present invention describes an isolated nucleotide fragment (I)

CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a

CC polypeptide (PP) having receptor-like protein kinase activity, mitogen

CC activated protein (MAP)-kinase activity, LIP15-like transcription factor

CC activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like

CC activity and CKC-like transcription factor activity. Also described: (1)

CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or

CC (II), operably linked to a regulatory sequence; (3) a plant (IV)

CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)

CC oil obtained from (V). (I) or its part can be used in antisense

CC inhibition or co-suppression in a transformed plant. (III) is useful for

CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,

CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for

CC creating transgenic plants having altered lipid profiles. (I) can also be

CC used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to

CC ABR40879 represent sequences used in the exemplification of the present

CC invention

XX Sequence 412 AA;

SQ

Query Match 90.6%; Score 29; DB 6; Length 412;

Best Local Similarity 80.0%; Pred. No. 2.8e+03;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5

Db 325 CRGDD 329

RESULT 41

ADI16354

ID ADI16354 standard; protein; 435 AA.

XX AC ADI16354;

XX DT 22-APR-2004 (first entry)

XX DE Human protein modification and maintenance molecule (PMMW) protein #39.

XX human; protein modification and maintenance molecule; PMMW;

KW gastrointestinal disorders; peptic ulcer; Crohn's disease;

KW cardiovascular disorders; hypertension; congenital heart disease;

KW autoimmune disease; inflammatory disease; AIDS; anaemia;

KW developmental disorder; Cushing's syndrome; tubular acidosis;

KW epithelial disorder; eczema; scabies; neurological disorder;

KW Alzheimer's disease; multiple sclerosis; infection; cancer.

XX Homo sapiens.

OS

XX WO2003100016-A2.

PN

XX 04-DEC-2003.

PD

XX 22-MAY-2003; 2003WO-US016498.

PF

XX 22-MAY-2002; 2002US-0383491P.

PR 24-JUN-2002; 2002US-0391378P.

PR 22-JUL-2002; 2002US-0397921P.

XX (INCY-) INCYTE CORP.

PA Khare R, Bulloch SA, Swarnakar A, Elliott VS, Marquis JP;

PI Mason PM, Chawla NK, Ramkumar J, Kable AE, Hafalia AJA, Lee SY;

PI Tran UK, Yue H, Becha SD, Griffin JA, Chang H, Jiang X, Jackson AA;

PI Richardson TW, Lal PG, Yao MG, Lu Y, Warren BA, Jin P, Wilson AD;

PI Gietzen KJ;

PI WPI; 2004-035124/03.

DR N-PSDB; ADI16406.

DR

XX New protein modification and maintenance molecules, useful for diagnosing

PT or treating e.g. peptic ulcer, hypertension, rheumatic fever, AIDS,

PT Cushing's syndrome, Alzheimer's disease, multiple sclerosis, stroke or

PT cancers.

XX Claim 1; SEQ ID NO 39; 419pp; English.

XX The invention comprises the amino acid and coding sequences of human

CC protein modification and maintenance molecules (PMMW). The DNA and

CC protein sequences of the invention are useful for the diagnosis and

CC treatment of disorders associated with expression of PMMW, such as:

CC gastrointestinal disorders (e.g. peptic ulcer and Crohn's disease),

CC cardiovascular disorders (e.g. hypertension and congenital heart
 CC disease), autoimmune or inflammatory disease (e.g. AIDS and anaemia),
 CC developmental disorders (e.g. Cushing's syndrome and tubular acidosis),
 CC epithelial disorders (e.g. eczema and scabies), neurological disorders
 CC (e.g. Alzheimer's disease and multiple sclerosis), infections and cancer.
 CC The present amino acid sequence represents a human PMW protein of the
 CC invention.

XX SQ Sequence 435 AA;

Query Match 90.6%; Score 29; DB 8; Length 435;

Best Local Similarity 80.0%; Pred. No. 3e+03;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5

Db 261 CKGDD 265

RESULT 42

ID ABB98131 standard; protein; 508 AA.

XX AC ABB98131;

XX DT 17-OCT-2002 (first entry)

XX DE Human PMW Incyte ID 7485421CD1.

XX Human; PMW; protein modification and maintenance molecule;
 KW anticonvulsant; neuroprotective; nootropic; cytostatic; antipsoriatic;
 KW antiaesthetic; dermatological; antidiabetic; antiparkinsonian;
 KW antianaemic; antiinflammatory; antiulcer; antitanginal; cardiant;
 KW hepatotropic; osteopathic; antiemetic; antipyretic; virucide;
 KW antibacterial; fungicide; gastrointestinal; antidiarrheic; laxative;
 KW haemostatic; anti-HIV; thrombolytic; anticoagulant; gynaecological;
 KW cardiovascular; antiarteriosclerotic; hypotensive; vasotropic;
 KW antitumour; antirheumatic; immunosuppressive; antiallergic; antithyroid;
 KW nephrotropic; antitout; thymimetic; antiarthritic; uteropathic;
 KW ophthalmological; antiparasitic; tranquiliser; vulnerary; keratolytic;
 KW auditory; antiseborrheic; antidepressant; neuroleptic; antiinfertility;
 KW anelminthic; protozoacide; Crohn's disease; hypertension; autoimmune;
 KW inflammatory; anaemia; cell proliferative; developmental; epithelial;
 KW scabies; neurological; Alzheimer's disease; reproductive;
 KW ectopic pregnancy; gene therapy; vaccine; disorder;
 KW ubiquitin specific protease; chromosome 1.

XX OS Homo sapiens.

XX PN W0200246383-A2.

XX PD 13-JUN-2002.

XX PF 05-DEC-2001; 2001WO-US046964.

XX PR 08-DEC-2000; 2000US-0254399P.

XX PR 21-DEC-2000; 2000US-0257803P.

XX PR 05-JAN-2001; 2001US-0260110P.

XX PR 19-JAN-2001; 2001US-0262851P.

XX PR 25-JAN-2001; 2001US-0264623P.

XX PA (INCY-) INCYTE GENOMICS INC.

XX PI Yue H, Azimzai Y, Kallick DA, Baughn MR, Griffin JA, Swarnakar A;

XX PI Lal PG, Walia NK, Hafalia AJA, Gandhi AR, Au-Young J, Elliott VS;

XX PI Ramkumar J, Thangavelu K, Lu Y, Warren BA, Lu DAM, Lee EA;

XX PI Tribouley CM, Arvizu C, Delegeane AM, Yao MG, Khan FA;

XX PI Sanjanwala MM;

XX DR WPI; 2002-519664/55.

XX DR N-PSDB; ABQ75952.

XX PT New isolated Protein Modification and Maintenance polypeptides, useful

PT for diagnosis, and treatment of e.g. gastrointestinal disorders.

XX Claim 1 (a); Page 167-168; 200pp; English.

XX The invention relates to an isolated Protein Modification and Maintenance
 CC (PMW) polypeptide. Polypeptides of the invention may be used in the
 CC diagnosis, treatment and prevention of disorders associated with
 CC decreased expression or activity of PMW. These include gastrointestinal
 CC disorders (e.g. Crohn's disease), cardiovascular disorders (e.g.
 CC hypertension), autoimmune/inflammatory disorders (e.g. anaemia), cell
 CC proliferative disorders, developmental disorders, epithelial disorders
 CC (e.g. scabies), neurological disorders (e.g. Alzheimer's disease)
 CC reproductive disorders (e.g. ectopic pregnancy), and in gene therapy or a
 CC vaccine for such diseases. They may also be used in the assessment of the
 CC effects of exogenous compound on the expression of nucleic acid and amino
 CC acid sequences of protein modification and maintenance molecules. The
 CC current sequence represents a human PMW of the invention, which has been
 CC found to have homology with mouse ubiquitin specific protease. The gene
 CC encoding this polypeptide has been localised to chromosome 1

XX SQ Sequence 508 AA;

Query Match 90.6%; Score 29; DB 5; Length 508;

Best Local Similarity 80.0%; Pred. No. 3.4e+03;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5

Db 261 CKGDD 265

RESULT 43

ADF22397

ID ADF92397 standard; protein; 514 AA.

XX AC ADF92397;

XX DT 26-FEB-2004 (first entry)

XX DE Human ubiquitin-specific protease KIAA1063.

XX ubiquitin-specific protease; Ub; USP; de-ubiquitination;

XX antiparkinsonian; nootropic; neuroprotective; neurodegenerative;

XX Alzheimer's disease; Parkinson's; muscular atrophy; gene therapy; human;

XX KIAA1063.

XX OS Homo sapiens.

XX PN JP2003189883-A.

XX PD 08-JUL-2003.

XX PF 30-SEP-2002; 2002JP-00287039.

XX PR 28-SEP-2001; 2001JP-00304709.

XX PA (DAUC) DAIICHI PHARM CO LTD.

XX WPI; 2003-856760/80.

XX DR N-PSDB; ADF92398.

XX Novel human ubiquitin specific protease having de-ubiquitination

PT activity, useful in diagnosis and treatment of muscular atrophy and

PT neurodegenerative diseases such as Alzheimer's disease, Parkinson's

PT disease.

XX Claim 1; SEQ ID NO 11; 117pp; Japanese.

XX The invention relates to a novel ubiquitin (Ub)-specific protease (USP)
 CC polypeptide having de-ubiquitination activity. The polypeptide of the
 CC invention demonstrates antiparkinsonian, nootropic, neuroprotective and
 CC muscular activities and may be useful for treating neurodegenerative
 CC diseases such as Alzheimer's disease, Parkinson's disease and muscular

CC atrophy, as well as in gene therapy applications. The current sequence is
XX that of the human ubiquitin-specific protease of the invention.

SQ Sequence 514 AA;

Query Match 90.6%; Score 29; DB 7; Length 514;
Best Local Similarity 80.0%; Pred. No. 3.5e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5

Db 267 CRGDD 271

RESULT 44

ABG10994

ID ABG10994 standard; protein; 525 AA.

XX AC ABG10994;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #10985.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR N-PSDB; AAS75181.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

XX PS Claim 20; SEQ ID NO 41353; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 525 AA;

Query Match 90.6%; Score 29; DB 4; Length 525;
Best Local Similarity 80.0%; Pred. No. 3.5e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5

Db 278 CRGDD 282

RESULT 45

ADH54836

ID ADH54836 standard; protein; 525 AA.

XX AC ADH54836;

XX DT 25-MAR-2004 (first entry)

XX DE Human ubiquitin protease 8 (hUBP8/Ubp22), SEQ ID NO:1.

XX KW Human; hUBP8; ubiquitin protease 8; ubiquitin specific protease 22;
KW Ubp22; USP22; Ubp8 homologue; deubiquitylation; histone H2B;
KW histone modification; transcriptional regulation; UBp8 modulator;
KW drug screening; tumorigenic potential; cancer cell sensitivity; cancer;
KW tumour; inflammation; viral infection; fungal infection;
KW Alzheimer's disease; human immunodeficiency virus; HIV infection;
KW hypercholesterolaemia; cytostatic; antiinflammatory; virucide; fungicide;
KW nootropic; neuroprotective; anti-HIV; antilipaemic; enzyme.

XX OS Saccharomyces cerevisiae.

XX PN WO2004003004-A1.

XX PD 08-JAN-2004.

XX PF 27-JUN-2003; 2003WO-US016238.

XX PR 27-JUN-2002; 2002US-0392335P.

XX PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.

XX PI Berger SL, Henry KW, McMahon SB;

XX WPI; 2004-083014/08.

XX PT A composition altering the expression of human ubiquitin protease 8, for
PT treating or diagnosing a disease associated with the expression of a gene
PT that relies upon deubiquitylation for transcription, e.g. inflammation or
PT cancer.

XX PS Claim 2; SEQ ID NO 1; 78pp; English.

XX CC The invention relates to a composition that alters the expression of
CC human ubiquitin protease 8 (hUBP8; also known as ubiquitin specific
CC protease 22, Ubp22 or USP22; ADH54836) or its homologue, and reduces,
CC eliminates or enhances its deubiquitylating activity in a cell. The
CC composition is preferably a double-stranded RNA with a sense and
CC antisense regions, or a protein ligand of hUBP8 such as an antibody or
CC fragment thereof, which are capable of inhibiting hUBP8 expression. hUBP8
CC is required for chromatin development and regulation of genes in the
CC histone-modifying SAGA complex, particularly histone H2B. Overexpression
CC of hUBP8 can result in overproduction of certain genes such as oncogenes,
CC while reduced hUBP8 enzymatic activity may reduce the production of
CC tumour suppressor proteins. The invention also relates to an expression
CC construct encoding hUBP8; methods of identifying modulators of hUBP8
CC activity or expression; a method of determining the tumorigenic potential
CC in a cell; a method of assessing the sensitivity of a tumour cell to an
CC agent that disrupts deubiquitylation; regarding the growth of a cancer
CC cell; and a method of treating or diagnosing a disease associated with
CC the expression of a gene that relies upon deubiquitylation for
CC transcription. The hUBP8 modulatory compositions of the invention are

CC useful for treating or diagnosing a disease associated with the
 CC expression of a gene that relies upon deubiquitylation for transcription,
 CC e.g. inflammation, viral infection, fungal infection, cancer, Alzheimer's
 CC disease, HIV infection or hypercholesterolaemia. The composition is also
 CC useful in rendering a cancer cell more sensitive to additional anti-
 CC tumour therapies, and in determining whether a cell is tumorigenic. The
 CC present sequence represents hUBP8/UBp22 which is referred to in the
 CC claims.

XX SQ Sequence 525 AA;

Query Match 90.6%; Score 29; DB 8; Length 525;
 Best Local Similarity 80.0%; Pred. No. 3.5e+03;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
 |:|
 Db 278 CKGDD 282

RESULT 46
 ABM81830
 ID ABM81830 standard; protein; 525 AA.

XX AC ABM81830;

XX DT 18-NOV-2004 (first entry)

XX DB Tumour-associated antigenic target (TAT) polypeptide PRO82683, SEQ:4711.

XX KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
 KW tumour; diagnosis; cell proliferative disorder; breast cancer;
 KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
 KW central nervous system cancer; bladder cancer; pancreatic cancer;
 KW cervical cancer; melanoma; leukaemia; hybridisation probe;
 KW chromosome identification; chromosome mapping; gene mapping;
 KW gene therapy; cytostatic.

XX OS Homo sapiens.

XX PN WO2004030615-A2.

XX PD 15-APR-2004.

XX PF 29-SEP-2003; 2003WO-US028547.

XX PR 02-OCT-2002; 2002US-0414971P.

XX PA (GETH) GENENTECH INC.

XX PI Wu TD, Zhang Z, Zhou Y;

XX DR WPI; 2004-347921/32.

XX DR N-PSDB; ACN40124.

XX PT New tumor-associated antigenic target polypeptides and nucleic acids,
 PT useful in preparing a medicament for treating or detecting a
 PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
 PT prostate cancer or tumor.

XX PS Claim 12; SEQ ID NO 4711; 7273pp; English.

XX CC The invention relates to human tumour-associated antigenic target (TAT)
 CC polypeptides, and their related nucleic acids. The TAT polypeptides are
 CC overexpressed in cancer tissues compared to normal tissues, and may thus
 CC serve as effective targets for the diagnosis and treatment of cancer in
 CC mammals. The invention also relates to nucleic acid and polypeptide
 CC sequences at least 80% identical to the TAT nucleic acids and
 CC polypeptides; expression vectors and host cells comprising a TAT nucleic
 CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
 CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
 CC TAT polypeptide; and methods and compositions for the treatment or
 CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,

CC antibodies, antagonists, binding molecules and compositions are useful
 CC for diagnosing or treating a cell proliferative disorder associated with
 CC increased TAT expression, particularly cancers such as breast cancer,
 CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
 CC cancer, pancreatic cancer, cervical cancer, cancers of the central
 CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
 CC used as hybridisation probes, in chromosome and gene mapping, in
 CC chromosome identification and in gene therapy. The present sequence
 CC represents a TAT polypeptide of the invention

XX SQ Sequence 525 AA;

Query Match 90.6%; Score 29; DB 8; Length 525;
 Best Local Similarity 80.0%; Pred. No. 3.5e+03;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
 |:|
 Db 278 CKGDD 282

RESULT 47

AAU23205

ID AAU23205 standard; protein; 528 AA.

XX AC AAU23205;

XX DT 18-DEC-2001 (first entry)

XX DE Novel human enzyme polypeptide #291.

XX KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
 KW ligase; hyperproliferative disorder; immunodeficiency disorder;
 KW autoimmune disorder; neurological disorder; metabolic disorder;
 KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
 KW blood-related disorder; infectious disorder; cytostatic; anti arthritic;
 KW nephrotropic; anticoagulant.

XX OS Homo sapiens.

XX PN WO200155301-A2.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001239.

XX PR 31-JAN-2000; 2000US-0179065P.

XX PR 04-FEB-2000; 2000US-0180628P.

XX PR 24-FEB-2000; 2000US-0184664P.

XX PR 02-MAR-2000; 2000US-0186350P.

XX PR 16-MAR-2000; 2000US-0189874P.

XX PR 17-MAR-2000; 2000US-0190076P.

XX PR 18-APR-2000; 2000US-0198123P.

XX PR 19-MAY-2000; 2000US-0205515P.

XX PR 07-JUN-2000; 2000US-0209467P.

XX PR 28-JUN-2000; 2000US-0214886P.

XX PR 30-JUN-2000; 2000US-0215135P.

XX PR 07-JUL-2000; 2000US-0216647P.

XX PR 07-JUL-2000; 2000US-0216800P.

XX PR 11-JUL-2000; 2000US-0217487P.

XX PR 11-JUL-2000; 2000US-0217496P.

XX PR 14-JUL-2000; 2000US-0218290P.

XX PR 26-JUL-2000; 2000US-0220963P.

XX PR 26-JUL-2000; 2000US-0220964P.

XX PR 14-AUG-2000; 2000US-0224518P.

XX PR 14-AUG-2000; 2000US-0224519P.

XX PR 14-AUG-2000; 2000US-0225213P.

XX PR 14-AUG-2000; 2000US-0225214P.

XX PR 14-AUG-2000; 2000US-0225266P.

XX PR 14-AUG-2000; 2000US-0225267P.

XX PR 14-AUG-2000; 2000US-0225268P.

XX PR 14-AUG-2000; 2000US-0225270P.

XX PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 23-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 29-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246533P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-465566/50.
DR N-PSDB; AAS41075.
XX
XX Novel polypeptides and polynucleotides useful for diagnosing, preventing,
PT treating neural, immune system, muscular, reproductive, pulmonary,
PT cardiovascular, renal, proliferative disorders and cancerous diseases.
XX
XX Claim 11; SEQ ID NO 1201; 1180pp; English.
XX
XX The present invention relates to the isolation of novel human enzyme
CC polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences
CC encoding them. The enzyme polypeptides of the invention may comprise the
CC functional classes of oxidoreductases, transferases, hydrolases, lyases,
CC isomerases or ligases. The sequences of the invention are useful in the
CC diagnosis, treatment, prevention and/or prognosis of a wide range of
CC disorders including hyperproliferative disorders (e.g. cancer),
CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g.
CC arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic
CC disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma),
CC cardiovascular disorders (e.g. atherosclerosis), blood-related disorders
CC (e.g. haemophilia), reproductive disorders (e.g. infertility) and
CC infectious disorders (e.g. Influenza). The polynucleotides of the
CC invention can also be used in gene therapy. AAU22915-AAU23814 represent
CC the novel human enzyme polypeptides of the invention. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 528 AA;
SQ
Query Match 90.6%; Score 29; DB 4; Length 528;
Best Local Similarity 80.0%; Pred. No. 3.5e+03;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
|:|:|
Db 281 CKGDD 285

RESULT 48
ADC10088
ID ADC10088 standard; protein; 556 AA.
XX
AC ADC10088;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human NOVX polypeptide SEQ ID NO: 108.
XX
KW cytostatic; antidiabetic; anorectic; cerebroprotective; neuroprotective;
KW antiinflammatory; gene therapy; antisense therapy; thymimetic; NOVX;
KW pathology; cancer; diabetes; obesity; endocrine disorder; CNS disorder;
KW inflammatory disorder; chromosome mapping; tissue typing;
KW predictive medicine.
XX
OS Homo sapiens.
XX
PN WO200300842-A2.
XX
PD 03-JAN-2003.
XX
PF 04-JUN-2002; 2002WO-US017443.
XX
PR 04-JUN-2001; 2001US-0295607P.
PR 04-JUN-2001; 2001US-0295661P.
PR 06-JUN-2001; 2001US-0296404P.
PR 06-JUN-2001; 2001US-0296418P.
PR 07-JUN-2001; 2001US-0296575P.
PR 11-JUN-2001; 2001US-0297414P.
PR 12-JUN-2001; 2001US-0295573P.
PR 12-JUN-2001; 2001US-0297567P.
PR 14-JUN-2001; 2001US-0298285P.
PR 15-JUN-2001; 2001US-0298528P.
PR 18-JUN-2001; 2001US-0299133P.
PR 19-JUN-2001; 2001US-0299499P.
PR 21-JUN-2001; 2001US-0300883P.
PR 22-JUN-2001; 2001US-0300177P.
PR 26-JUN-2001; 2001US-0300883P.
PR 28-JUN-2001; 2001US-0301530P.
PR 28-JUN-2001; 2001US-0301550P.
PR 03-JUL-2001; 2001US-0302951P.
PR 31-JUL-2001; 2001US-0308890P.
PR 14-SEP-2001; 2001US-0322297P.
PR 25-SEP-2001; 2001US-0324659P.
PR 03-DEC-2001; 2001US-0337477P.
PR 14-DEC-2001; 2001US-0341562P.
PR 21-FEB-2002; 2002US-0358565P.
PR 21-FEB-2002; 2002US-0359122P.
PR 22-FEB-2002; 2002US-0358978P.
PR 22-FEB-2002; 2002US-0359034P.
PR 22-FEB-2002; 2002US-0359035P.
PR 22-FEB-2002; 2002US-0359121P.
PR 27-FEB-2002; 2002US-0359964P.
PR 01-MAR-2002; 2002US-0360858P.
PR 12-MAR-2002; 2002US-0363430P.
PR 12-MAR-2002; 2002US-0363676P.
PR 10-APR-2002; 2002US-0371346P.
PR 10-MAY-2002; 2002US-0379444P.
PR 04-JUN-2002; 2002US-00379444.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Agee ML, Anderson DW, Berghs C, Casman SJ, Catterton E;
PI Dipippo VA, Edinger SR, Eisen A, Ellerman K, Gangolli EA;
PI Gerlach VL, Gorman L, Guo X, Herrmann JL, Hjalt I, Ji W, Kekuda R;

PI Khramtsov NV, Li L, Liu X, Malyankar UM, Miller CE, Millet I;
PI Ort T, Padigar M, Patturajan M, Pena CEA, Rastelli L, Rieger DK;
PI Rothenberg ME, Shenoy SG, Shinkets RA, Smithson G, Spaderna SK;
PI Spytek KA, Stone DJ, Vernet CAM, Zhong H, Zhong M, Alsobrook JP;
PI Burgess CE, Lepley DM;
XX
DR WPI; 2003-210149/20.
DR N-PSDB; ADC10087.
XX
XX
PT New isolated NOVX polypeptides and nucleic acid molecules useful for
PT treating, preventing and diagnosing pathological conditions with NOVX-
PT associated disorders, such as cancer, obesity, diabetes and inflammatory
PT or CNS diseases.
XX
PS Claim 1; SEQ ID NO 108; 772pp; English.
XX
CC The invention relates to novel isolated polypeptides, mature form of the
CC polypeptide, a sequence that is 95% identical to the polypeptide or the
CC polypeptide comprising one or more conservative substitutions. The NOVX
CC polypeptide is useful for treating or preventing a pathology associated
CC with the polypeptide e.g. disorders associated with aberrant expression
CC or activity of the polypeptide, such as cancer, diabetes, obesity, and
CC endocrine, CNS and inflammatory disorders. They can also be used in
CC various detection and screening assays, chromosome mapping, tissue typing
CC and predictive medicine. This sequence corresponds to one of the
CC polypeptides of the invention.
XX
SQ Sequence 556 AA;
Query Match 90.6%; Score 29; DB 7; Length 556;
Best Local Similarity 80.0%; Pred. No. 3.7e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
|:|:|
Db 309 CKGDD 313

RESULT 49
ADC10090
ID ADC10090 standard; protein; 575 AA.
XX
AC ADC10090;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human NOVX polypeptide SEQ ID NO: 110.
XX
KW cytostatic; antidiabetic; anorectic; cerebroprotective; neuroprotective;
KW antiinflammatory; gene therapy; antisense therapy; thymimetic; NOVX;
KW pathology; cancer; diabetes; obesity; endocrine disorder; CNS disorder;
KW inflammatory disorder; chromosome mapping; tissue typing;
KW predictive medicine.
XX
OS Homo sapiens.
XX
PN WO200300842-A2.
XX
PD 03-JAN-2003.
XX
PF 04-JUN-2002; 2002WO-US017443.
XX
PR 04-JUN-2001; 2001US-0295607P.
PR 04-JUN-2001; 2001US-0295661P.
PR 06-JUN-2001; 2001US-0296404P.
PR 06-JUN-2001; 2001US-0296418P.
PR 07-JUN-2001; 2001US-0296575P.
PR 11-JUN-2001; 2001US-0297414P.
PR 12-JUN-2001; 2001US-0295573P.
PR 12-JUN-2001; 2001US-0297567P.
PR 14-JUN-2001; 2001US-0298285P.
PR 15-JUN-2001; 2001US-0298528P.
PR 18-JUN-2001; 2001US-0299133P.

PR 19-JUN-2001; 2001US-0299230P.
PR 21-JUN-2001; 2001US-0299949P.
PR 22-JUN-2001; 2001US-0300177P.
PR 26-JUN-2001; 2001US-0300883P.
PR 28-JUN-2001; 2001US-0301530P.
PR 28-JUN-2001; 2001US-0301550P.
PR 03-JUL-2001; 2001US-0302951P.
PR 31-JUL-2001; 2001US-0308850P.
PR 14-SEP-2001; 2001US-0322297P.
PR 25-SEP-2001; 2001US-0324669P.
PR 03-DEC-2001; 2001US-0337477P.
PR 14-DEC-2001; 2001US-0341562P.
PR 21-FEB-2002; 2002US-0358656P.
PR 21-FEB-2002; 2002US-0359122P.
PR 22-FEB-2002; 2002US-0358978P.
PR 22-FEB-2002; 2002US-0359034P.
PR 22-FEB-2002; 2002US-0359035P.
PR 22-FEB-2002; 2002US-0359121P.
PR 27-FEB-2002; 2002US-0359964P.
PR 01-MAR-2002; 2002US-0360858P.
PR 12-MAR-2002; 2002US-0363430P.
PR 12-MAR-2002; 2002US-0363676P.
PR 10-APR-2002; 2002US-0371346P.
PR 10-MAY-2002; 2002US-0379444P.
PR 04-JUN-2002; 2002US-00379444.

(CURA-) CURAGEN CORP.

XX Agee ML, Anderson DW, Berghs C, Casman SJ, Catterton E;
PI Dipippo VA, Edinger SR, Eissen A, Ellerman K, Gangolli EA;
PI Gerlach VL, Gorman L, Guo X, Herrmann JM, Hjalt T, Ji W, Kekuda R;
PI Khramtsov NV, Li L, Liu X, Malyankar UM, Miller CE, Millet I;
PI Ort T, Padigar M, Patturajan M, Pena CEA, Rastelli L, Rieger DK;
PI Rothenberg ME, Shenoy SG, Shinkets RA, Smithson G, Spaderna SK;
PI Spytek KA, Stone DJ, Vernet CAM, Zhong H, Zhong M, Alsobrook JP;
PI Burgess CE, Lepley DM;

XX WPI: 2003-210149/20.
DR N-PSDB; ADC10089.
XX

XX New isolated NOVX polypeptides and nucleic acid molecules useful for
PT treating, preventing and diagnosing pathological conditions with NOVX-
PT associated disorders, such as cancer, obesity, diabetes and inflammatory
PT or CNS diseases.

XX Claim 1; SEQ ID NO 110; 772pp; English.

XX The invention relates to novel isolated polypeptides, mature form of the
CC polypeptide, a sequence that is 95% identical to the polypeptide or the
CC polypeptide comprising one or more conservative substitutions. The NOVX
CC polypeptide is useful for treating or preventing a pathology associated
CC with the polypeptide e.g. disorders associated with aberrant expression
CC or activity of the polypeptide, such as cancer, diabetes, obesity, and
CC endocrine, CNS and inflammatory disorders. They can also be used in
CC various detection and screening assays, chromosome mapping, tissue typing
CC and predictive medicine. This sequence corresponds to one of the
CC polypeptides of the invention.

XX Sequence 575 AA;

Query Match 90.6%; Score 29; DB 7; Length 575;

Best Local Similarity 80.0%; Pred. No. 3.8e+03;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5

[:|||]

Db 328 CRGDD 332

RESULT 50

ADC10086

ID ADC10086 standard; protein; 581 AA.

XX

AC ADC10086;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human NOVX polypeptide SEQ ID NO: 106.
XX

KW cytostatic; antidiabetic; anorectic; cerebrotective; neuroprotective;
KW antiinflammatory; gene therapy; antisense therapy; thyromimetic; NOVX;
KW pathology; cancer; diabetes; obesity; endocrine disorder; CNS disorder;
KW inflammatory disorder; chromosome mapping; tissue typing;
KW predictive medicine.

XX Homo sapiens.

XX WO2003000842-A2.

XX 03-JAN-2003.

XX 04-JUN-2002; 2002WO-US017443.

XX 04-JUN-2001; 2001US-0295607P.

XX 04-JUN-2001; 2001US-0295661P.

XX 06-JUN-2001; 2001US-0296404P.

XX 06-JUN-2001; 2001US-0296418P.

XX 07-JUN-2001; 2001US-0296575P.

XX 11-JUN-2001; 2001US-0297414P.

XX 12-JUN-2001; 2001US-0295573P.

XX 12-JUN-2001; 2001US-0297567P.

XX 14-JUN-2001; 2001US-0298285P.

XX 15-JUN-2001; 2001US-0298528P.

XX 18-JUN-2001; 2001US-0299133P.

XX 19-JUN-2001; 2001US-0299230P.

XX 21-JUN-2001; 2001US-0299949P.

XX 22-JUN-2001; 2001US-0300177P.

XX 26-JUN-2001; 2001US-0300883P.

XX 28-JUN-2001; 2001US-0301530P.

XX 28-JUN-2001; 2001US-0301550P.

XX 03-JUL-2001; 2001US-0302951P.

XX 31-JUL-2001; 2001US-0308900P.

XX 14-SEP-2001; 2001US-0322297P.

XX 25-SEP-2001; 2001US-0324669P.

XX 03-DEC-2001; 2001US-0337477P.

XX 14-DEC-2001; 2001US-0341562P.

XX 21-FEB-2002; 2002US-0358656P.

XX 21-FEB-2002; 2002US-0359122P.

XX 22-FEB-2002; 2002US-0358978P.

XX 22-FEB-2002; 2002US-0359034P.

XX 22-FEB-2002; 2002US-0359035P.

XX 22-FEB-2002; 2002US-0359121P.

XX 27-FEB-2002; 2002US-0359964P.

XX 01-MAR-2002; 2002US-0360858P.

XX 12-MAR-2002; 2002US-0363430P.

XX 12-MAR-2002; 2002US-0363676P.

XX 10-APR-2002; 2002US-0371346P.

XX 10-MAY-2002; 2002US-0379444P.

XX 04-JUN-2002; 2002US-00379444.

(CURA-) CURAGEN CORP.

XX Agee ML, Anderson DW, Berghs C, Casman SJ, Catterton E;
PI Dipippo VA, Edinger SR, Eissen A, Ellerman K, Gangolli EA;
PI Gerlach VL, Gorman L, Guo X, Herrmann JM, Hjalt T, Ji W, Kekuda R;
PI Khramtsov NV, Li L, Liu X, Malyankar UM, Miller CE, Millet I;
PI Ort T, Padigar M, Patturajan M, Pena CEA, Rastelli L, Rieger DK;
PI Rothenberg ME, Shenoy SG, Shinkets RA, Smithson G, Spaderna SK;
PI Spytek KA, Stone DJ, Vernet CAM, Zhong H, Zhong M, Alsobrook JP;
PI Burgess CE, Lepley DM;

XX WPI: 2003-210149/20.

XX N-PSDB; ADC10085.

XX New isolated NOVX polypeptides and nucleic acid molecules useful for
PT treating, preventing and diagnosing pathological conditions with NOVX-

PT associated disorders, such as cancer, obesity, diabetes and inflammatory
 FT or CNS diseases.

PS Claim 1; SEQ ID NO 106; 772pp; English.

XX The invention relates to novel isolated polypeptides, mature form of the
 CC polypeptide, a sequence that is 95% identical to the polypeptide or the
 CC polypeptide comprising one or more conservative substitutions. The NOVX
 CC polypeptide is useful for treating or preventing a pathology associated
 CC with the polypeptide e.g. disorders associated with aberrant expression
 CC or activity of the polypeptide, such as cancer, diabetes, obesity, and
 CC endocrine, CNS and inflammatory disorders. They can also be used in
 CC various detection and screening assays, chromosome mapping, tissue typing
 CC and predictive medicine. This sequence corresponds to one of the
 CC polypeptides of the invention.

XX Sequence 581 AA;

Query Match 90.6%; Score 29; DB 7; Length 581;
 Best Local Similarity 80.0%; Pred. No. 3.9e+03;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5

DB 334 CKGDD 338
 : : : :
 : : : :
 : : : :

RESULT 51

AAE17560
 ID AAE17560 standard; protein; 593 AA.

AC AAE17560;

DT 22-APR-2002 (first entry)

DE Human pancreatic tumour protein, KIAA1063.

XX Human; pancreatic tumour protein; pancreatic cancer; gene therapy;
 KW vaccine.

OS Homo sapiens.

PN W0200194409-A2.

XX 13-DEC-2001.

XX 31-MAY-2001; 2001WO-US018003.

XX 07-JUN-2000; 2000US-0210329P.

PR 14-NOV-2000; 2000US-0248980P.

PR 15-MAY-2001; 2001US-0291197P.

XX (CORI-) CORIXA CORP.

XX First SK, Harlocker SL, Dillon DC, Kalos MD;

DR WPI; 2002-154565/20.

DR N-PSDB; AAD28257.

XX Pancreatic tumor polypeptide and polynucleotide useful in diagnosis,
 PT prevention and/or treatment of cancer, especially pancreatic cancer.

PS Claim 2a; Page 107-108; 115pp; English.

XX The invention relates to human pancreatic tumour polypeptides and nucleic
 CC acid molecules encoding such polypeptides. The invention also relates to
 CC compositions and methods for the diagnosis, prevention and therapy of
 CC cancer, particularly pancreatic cancer. Sequences of the invention are
 CC used in gene therapy and in vaccines. The present sequence is human
 CC pancreatic tumour protein

XX Sequence 593 AA;

Query Match 90.6%; Score 29; DB 5; Length 593;
 Best Local Similarity 80.0%; Pred. No. 3.9e+03;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5

DB 346 CKGDD 350
 : : : :
 : : : :
 : : : :

RESULT 52

ADH54838

ID ADH54838 standard; protein; 593 AA.

XX ADH54838;

XX 25-MAR-2004 (first entry)

XX Human KIAA1063 protein related to hUBP8/Ubp22, SEQ ID NO:3.

XX Human; KIAA1063 protein; hUBP8; ubiquitin protease 8;

KW ubiquitin specific protease 22; Ubp22; USP22; Ubp8 homologue;

KW deubiquitylation; histone H2B; histone modification;

KW transcriptional regulation; UBP8 modulator; drug screening;

KW tumorigenic potential; cancer cell sensitivity; cancer; tumour;

KW inflammation; viral infection; fungal infection; Alzheimer's disease;

KW human immunodeficiency virus; HIV infection; hypercholesterolemia;

KW cytostatic; antiinflammatory; virucide; fungicide; neurotropic;

KW neuroprotective; anti-HIV; antilipaemic; enzyme.

XX Saccharomyces cerevisiae.

OS Key Location/Qualifiers

FT Misc-difference 1..68

FT /note= "This sequence is not present in SEQ ID NO:1"

FT Misc-difference 69..593

FT /note= "Identical to SEQ ID NO:1"

XX W02004003004-A1.

XX 08-JAN-2004.

XX 27-JUN-2003; 2003WO-US016238.

XX 27-JUN-2002; 2002US-0392335P.

XX (WIST-) WISTAR INST ANATOMY & BIOLOGY.

XX Berger SL, Henry KW, McMahon SB;

DR WPI; 2004-083014/08.

DR GENBANK; BAA83015.

XX A composition altering the expression of human ubiquitin protease 8, for
 PT treating or diagnosing a disease associated with the expression of a gene
 PT that relies upon deubiquitylation for transcription, e.g. inflammation or
 PT cancer.

PS Disclosure; SEQ ID NO 3; 78pp; English.

XX The invention relates to a composition that alters the expression of
 CC human ubiquitin protease 8 (hUBP8; also known as ubiquitin specific
 CC protease 22, Ubp22 or USP22; ADH54836) or its homologue, and reduces,
 CC eliminates or enhances its deubiquitylating activity in a cell. The
 CC composition is preferably a double-stranded RNA with a sense and
 CC antisense regions, or a protein ligand of hUBP8 such as an antibody or
 CC fragment thereof, which are capable of inhibiting hUBP8 expression. hUBP8
 CC is required for chromatin development and regulation of genes in the
 CC histone-modifying SAGA complex, particularly histone H2B. Overexpression
 CC of hUBP8 can result in overproduction of certain genes such as oncogenes,
 CC while reduced hUBP8 enzymatic activity may reduce the production of
 CC tumour suppressor proteins. The invention also relates to an expression
 CC construct encoding hUBP8; methods of identifying modulators of hUBP8
 CC activity or expression; a method of determining the tumorigenic potential

CC in a cell; a method of assessing the sensitivity of a tumour cell to an
 CC agent that disrupts deubiquitylation; retarding the growth of a cancer
 CC cell; and a method of treating or diagnosing a disease associated with
 CC the expression of a gene that relies upon deubiquitylation for
 CC transcription. The hUPB8 modulatory compositions of the invention are
 CC useful for treating or diagnosing a disease associated with the
 CC expression of a gene that relies upon deubiquitylation for transcription,
 CC e.g. inflammation, viral infection, fungal infection, cancer, Alzheimer's
 CC disease, HIV infection or hypercholesterolaemia. The composition is also
 CC useful in rendering a cancer cell more sensitive to additional anti-
 CC tumour therapies, and in determining whether a cell is tumorigenic. The
 CC present sequence represents human KIAA1063 protein, which consists of the
 CC hUPB8/UBP22 amino acid sequence (ADH54836), plus an additional 68 N-
 CC terminal region.

XX SQ Sequence 593 AA;

Query Match 90.6%; Score 29; DB 8; Length 593;
 Best Local Similarity 80.0%; Pred. No. 3.9e+03;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
 Db 346 CRGDD 350

RESULT 53

ADQ20907 ADQ20907 standard; protein; 593 AA.

XX AC ADQ20907;

XX DT 26-AUG-2004 (first entry)

XX DE Human soft tissue sarcoma-upregulated protein - SEQ ID 3727.

XX KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.

XX OS Homo sapiens.

XX PN W02004048938-A2.

XX PD 10-JUN-2004.

XX PF 26-NOV-2003; 2003WO-US038193.

XX PR 26-NOV-2002; 2002US-0429739P.

XX PA (PROT-) PROTEIN DESIGN LABS INC.

XX PI Aziz N, Ginsburg WM, Zlotnik A;

XX DR WPI; 2004-441208/41.

XX Early detection of soft tissue sarcoma comprises determining expression
 PT of a gene in a first soft tissue sample and a normal soft tissue sample
 PT and comparing the gene expression, also useful in treating soft tissue
 PT sarcoma.

XX PS Example 2; SEQ ID NO 3727; 210pp; English.

XX The invention relates to a novel method for detecting soft tissue sarcoma
 CC which comprises obtaining a first soft tissue sample from an individual
 CC and a normal soft tissue sample from the same or different individual,
 CC determining the expression of a gene in both samples and comparing the
 CC of protein expression in the first soft tissue samples, where a higher level
 CC presence of soft tissue sarcoma. The method of the invention has
 CC cytostatic applications and may be useful for detecting soft tissue
 CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
 CC acid sequences may be useful in diagnostic and screening applications.
 CC The current sequence is that of a human soft tissue sarcoma-upregulated
 CC protein of the invention. The current sequence is not shown within the

CC specification per se but was submitted in CD format by the inventor.
 XX SQ Sequence 593 AA;

Query Match 90.6%; Score 29; DB 8; Length 593;
 Best Local Similarity 80.0%; Pred. No. 3.9e+03;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
 Db 346 CRGDD 350

RESULT 54

ABO70157 ABO70157 standard; protein; 639 AA.

XX AC ABO70157;

XX DT 29-JUL-2004 (first entry)

XX DE Pseudomonas aeruginosa polypeptide #2332.

XX KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX OS Pseudomonas aeruginosa.

XX PN US6551795-B1.

XX PD 22-APR-2003.

XX PF 18-FEB-1999; 99US-00252991.

XX PR 18-FEB-1998; 98US-0074788P.

XX PR 27-JUL-1998; 98US-0094150P.

XX PA (GENO-) GENOME THERAPEUTICS CORP.

XX PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX DR WPI; 2003-615309/58.

XX DR N-PSDB; ABD03728.

XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of
 PT pathological conditions resulting from bacterial infection.

XX PS Disclosure; SEQ ID NO 18903; 455pp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the
 CC polynucleotides encoding them. The sequences are useful in diagnosis and
 CC therapy of pathological conditions, as molecular targets for diagnostics,
 CC prophylaxis and treatment of pathological conditions resulting from a
 CC bacterial infection, for evaluating a compound, such as a polypeptide,
 CC for the ability to bind a P. aeruginosa nucleic acid, as components of
 CC effective antibacterial targets, as targets for antibacterial drugs,
 CC including anti-P. aeruginosa drugs, as templates for recombinant
 CC production of P. aeruginosa-derived peptides or polypeptides, as target
 CC components for diagnosis and/or treatment of P. aeruginosa-caused
 CC infection, and in detection of P. aeruginosa sequences or other sequences
 CC of Pseudomonas species using biochip technology. Sequences ABO67826-
 CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html

XX SQ Sequence 639 AA;

Query Match 90.6%; Score 29; DB 7; Length 639;
 Best Local Similarity 80.0%; Pred. No. 4.2e+03;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5

Db 602 CKGDD 606

RESULT 55
ABU92027
ID ABU92027 standard; protein; 936 AA.
XX
AC ABU92027;
XX
DT 15-JUL-2003 (first entry)
XX
DE Human protein modification and maintenance molecule-7 (PMMM-7).
XX
KW Human; protein modification and maintenance molecule; PMMM; cancer;
KW cell proliferation disorder; atherosclerosis; neurological disorder;
KW epilepsy; Huntington's disease; stroke; immune disorder; allergy;
KW inflammatory disorder; AIDS; developmental disorder; hypothyroidism;
KW Cushing's syndrome; gastrointestinal disorder; epithelial disorder;
KW infection; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic;
KW neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnerary;
KW antiinflammatory; thymomimetic.
XX
OS Homo sapiens.
XX
PN WO2003031939-A2.
XX
PD 17-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032850.
XX
PR 12-OCT-2001; 2001US-0329689P.
XX
PR 25-OCT-2001; 2001US-0335703P.
XX
PR 09-NOV-2001; 2001US-0348877P.
XX
PR 28-NOV-2001; 2001US-0334145P.
XX
PR 06-DEC-2001; 2001US-0337451P.
XX
PR 14-DEC-2001; 2001US-0340584P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Rankumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY;
PI Tran UK, Becha SD, Duggan BM, Lee EA, Griffin JA, Li JX;
PI Sprague JW, Hafalia AJA, Chawla NK, Lehr-Mason PM, Kable AE, Yue H;
PI Marquis JP, Yao MG, Richardson RW, Tang TY, Jin P, Chien D;
PI Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;
XX
WPI; 2003-430274/40.
DR N-PSDB; ACA92422.
XX
XX New human protein modification and maintenance molecules (PMMM), useful
PT for diagnosing, treating and preventing diseases or conditions associated
PT with the aberrant PMMM expression e.g. cancer, atherosclerosis, or
PT infections.
XX
PS Claim 1; Page 244-246; 311pp; English.
XX
CC The present invention relates to the isolation of human protein
CC modification and maintenance molecules (PMMM), and the polynucleotide
CC sequences encoding them. A total of 40 PMMM polypeptides (designated PMMM
CC -1 to PMMM-40) are disclosed. The sequences of the invention are useful
CC for diagnosing a condition or disease associated with the expression of
CC PMMM in a subject, preparing a polyclonal or monoclonal antibody, and
CC generating an expression profile of a sample containing the
CC polynucleotides. The diseases or conditions associated with decreased
CC expression or overexpression of PMMM are cell proliferation disorders
CC (e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,
CC Huntington's disease, stroke), immune/inflammatory disorders (e.g. AIDS,
CC allergies), developmental disorders (e.g. hypothyroidism, Cushing's
CC syndrome), gastrointestinal or epithelial disorders, and infections. The
CC PMMM polypeptides or their fragments are useful in screening compounds
CC for effectiveness as agonists or antagonists of the polypeptides, or in
CC altering the expression of the target polynucleotide and compounds that
CC specifically bind to, or modulate the activity of the polypeptide.

CC ABU92021-ABU92060 represent the human PMMM polypeptides of the invention
XX
SQ Sequence 936 AA;
Query Match 90.6%; Score 29; DB 6; Length 936;
Best Local Similarity 80.0%; Pred. No. 6e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 CRGDD 5
DB 678 CKGDD 682

RESULT 56
AAU82708
ID AAU82708 standard; protein; 953 AA.
XX
AC AAU82708;
XX
DT 23-APR-2002 (first entry)
XX
DE Amino acid sequence of novel human protease #7.
XX
KW Human; protease; cancer; immune-related disorder; cardiovascular disease;
KW neuronal-associated disease; metabolic disorder; inflammatory disorder;
KW nervous system disorder; sexual dysfunction; pain; mood disorder;
KW hypertension; psychotic disorder; neurological disorder; dyskinesia;
KW viral infection; human immunodeficiency virus; HIV; non-viral infection;
KW ocular disease; cytostatic; enzyme.
XX
OS Homo sapiens.
XX
PN WO200200860-A2.
XX
PD 03-JAN-2002.
XX
PF 26-JUN-2001; 2001WO-US020171.
XX
PR 26-JUN-2000; 2000US-0214047P.
XX
PA (SUGE-) SUGEN INC.
XX
PI Plowman G, Whyte D, Sudarsanam S, Manning G, Caenepeel S;
PI Charydczak G;
XX
WPI; 2002-139913/18.
DR N-PSDB; ABK31750.
XX
XX Nucleic acids encoding novel human proteases, useful for useful for
PT treating diseases and disorders such as cancers, immune-related diseases
PT and disorders, cardiovascular disease (e.g. restenosis) and inflammatory
PT disorders.
XX
PS Claim 6; Fig 2C-D; 313pp; English.
XX
CC The present invention relates to the isolation of novel human proteases,
CC and the nucleic acids encoding them. The sequences of the invention are
CC useful for treating diseases and disorders such as cancers (e.g. breast,
CC colon, lung), immune-related diseases and disorders (e.g. inflammatory
CC diseases and asthma), cardiovascular diseases (e.g. restenosis and
CC coronary thrombosis), brain or neuronal-associated diseases, metabolic
CC disorders (e.g. diabetes, obesity), inflammatory disorders (e.g.
CC rheumatoid arthritis and psoriasis), central or peripheral nervous system
CC diseases, migraines, pain, sexual dysfunction, mood disorders, attention
CC disorders, cognitive disorders, hypotension, hypertension, psychotic
CC disease) and dyskinesias. The nucleic acids and polypeptides are also
CC useful for treating viral infections caused by human immunodeficiency
CC virus (HIV), and non-viral infections such as ocular disease (e.g.
CC glaucoma) and macular degeneration. AAU82702-AAU82760 represent the novel
CC human proteases of the invention
XX
SQ Sequence 953 AA;

Query Match 90.6%; Score 29; DB 5; Length 953;
Best Local Similarity 80.0%; Pred. No. 6.1e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
|:|:|
Db 695 CKGDD 699

RESULT 57
AAU74761
ID AAU74761 standard; protein; 953 AA.

AC AAU74761;

XX 09-APR-2002 (first entry)

DT Human protease PRTS-21 protein sequence.

DE Human; protease; PRTS; gastrointestinal; Crohn's disease; cancer;
KW cardiovascular; atherosclerosis; autoimmune deficiency disorder; dermatitis;
KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;
KW cell proliferative disorder; developmental disorder; epilepsy;
KW Duchenne muscular dystrophy; epithelial disorder; neurological disorder;
KW reproductive disorder; endometriosis.

XX Homo sapiens.

XX WO200198468-A2.

XX 27-DEC-2001.

XX 13-JUN-2001; 2001WO-US019178.

XX 16-JUN-2000; 2000US-0212336P.

XX 22-JUN-2000; 2000US-0213955P.

XX 29-JUN-2000; 2000US-0215396P.

XX 07-JUL-2000; 2000US-0216821P.

XX 14-JUL-2000; 2000US-0218946P.

XX (INCY-) INCYTE GENOMICS INC.

XX Yue H, Elliott VS, Gandhi AR, Lal P, Au-Young J, Triboulevy CM;

XX Deleageane AM, Baughn MR, Nguyen DB, Lee EA, Hafalia A, Khan FA;

XX Walia NK, Yao MG, Lu DAM, Patterson C, Tang YT, Walsh RT;

XX Azimzai Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;

XX Kallick DA;

XX WPI; 2002-090437/12.

XX N-PSDB; ABK12904.

XX Twenty one human proteases (referred to as PRTS-1 to PRTS-21), useful in
PT the diagnosis, treatment and prevention of gastrointestinal (e.g.
PT gastritis), cardiovascular (e.g. atherosclerosis) and cell proliferative
PT (e.g. cancer) disorders.

XX Claim 1; Page 158-160; 177pp; English.

XX The present invention relates to twenty one new human proteases, referred
CC to as PRTS-1 to PRTS-21. The PRTS polynucleotides and polypeptides of the
CC invention are useful in the diagnosis, treatment and prevention of
CC gastrointestinal e.g. gastritis, esophageal carcinoma and Crohn's
CC disease, cardiovascular e.g. atherosclerosis, hypertension and myocardial
CC infarction, autoimmune/inflammatory e.g. acquired immunodeficiency
CC syndrome (AIDS), allergies and osteoarthritis, cell proliferative e.g.
CC cancer, developmental e.g. Duchenne and Becker muscular dystrophy,
CC epithelial e.g. dermatitis, neurological e.g. epilepsy and Alzheimer's
CC disease and reproductive e.g. infertility and endometriosis disorders.
CC Numerous other examples of each disorder are given in the specification.

XX The present protein sequence represents the human protease PRTS-21
XX protein of the invention

SQ Sequence 953 AA;

Query Match 90.6%; Score 29; DB 5; Length 953;
Best Local Similarity 80.0%; Pred. No. 6.1e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
|:|:|
Db 695 CKGDD 699

RESULT 58
AAR96415

ID AAR96415 standard; peptide; 9 AA.

XX AAR96415;

XX 21-OCT-2004 (revised)

DT 05-JUL-1996 (first entry)

XX RGD cyclic peptide, TL#753, binds alpha-v,beta3 integrin receptor.

DE RGD-containing peptide; alpha-v, beta-3 integrin receptor; osteoclast;
KW matrix; bone; inhibition; bone resorption; promote; endothelial cell;
KW smooth muscle cell; restenosis; angiogenesis; cyclic.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1

FT Disulfide-bond 2..8 /note= "Acyl-Ala"

FT Modified-site 9 /note= "Amidated C-terminal"

XX WO9528426-A2.

XX 26-OCT-1995.

XX 12-APR-1995; 95WO-US004741.

XX 13-APR-1994; 94US-00227316.

XX 08-SEP-1994; 94US-00303052.

XX (LJOL-) LA JOLLA CANCER RES FOUND.

XX Cheng S, Ingram R, Mullen D, Tschopp J;

XX WPI; 1995-373767/48.

XX Altering alpha-v, beta-3 integrin receptor-mediated binding of cell to
PT matrix - using conformationally restrained peptide of RGD type, e.g. for
PT treating inappropriate angiogenesis or for inhibiting bone resorption.

XX Claim 85; Page 80; 99pp; English.

XX The sequences given in AAR96301-417 are non-naturally occurring RGD-
CC containing peptides which alter the alpha-v, beta-3 integrin receptor
CC binding of a cell to a matrix, such as the binding of an osteoclast to a
CC matrix such as bone. These peptides inhibit bone resorption and can
CC inhibit or promote alpha-v, beta-3-mediated cell attachment depending on
CC whether they are present in the cell in a soluble form or are bound to a
CC solid substrate. These peptides can be used in the amelioration of the
CC severity of a pathology involving alpha-v, beta-3 receptor-mediated
CC binding of a cell, such as an osteoclast, endothelial cell or smooth
CC muscle cell to a matrix. They are used for treating conditions associated
CC with restenosis or inappropriate or insufficient angiogenesis, or for
CC inhibiting osteoclast binding to the matrix

XX Revised record issued on 21-OCT-2004 : Correction to Feature Table Key

XX Sequence 9 AA;

SQ

Query Match 87.5%; Score 28; DB 2; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.8e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 2 CRGDE 6

RESULT 59

AAW66841
ID AAW66841 standard; peptide; 9 AA.

XX AC AAW66841;

XX 25-MAR-2003 (revised)
DT 10-DEC-1998 (first entry)

XX Peptide useful for altering bone resorption.

XX bone resorption; pharmacore; angiogenesis; restenosis; integrin receptor;
KW alpha v beta 3 integrin receptor; osteoclast; cyclic.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 9 /note= "C-terminal amide"

XX US5807819-A.
PN

XX 15-SEP-1998.
PD

XX 12-APR-1995; 95US-00421698.
PF

XX 15-APR-1994; 94US-00227316.
PR

XX 08-SEP-1994; 94US-00303052.
PR

XX (LJOL-) LA JOLLA CANCER RES CENT.

XX Tachopp JP, Mullen D, Ingram R, Cheng S;
PI

XX WPI; 1998-555601/47.
DR

XX Use of peptide derivatives which can alter integrin receptor binding -
PT for altering bone resorption, treating angiogenesis or restenosis and
PT altering integrin receptor mediated interactions.

XX Example 2; Fig 2E; 87pp; English.
PS

XX A new method is claimed for altering bone resorption. It comprises
XX administration of a peptide of formula: X1X2X3X4GX5X6X7X8 ; where X1 =
XX R1R2N or 0-10 amino acids (optionally protected by acetylation at the N-
XX terminus); X2 = absent or 1 amino acid; X3 = absent or 1 or 2 amino acids
XX; X4 = N-Me-Arg; X5 = residue which provides an ionic interaction with an
XX integrin receptor, or is Msa, Paa or Tfsa; X6 = residue which has an
XX aliphatic side chain; a non-natural amino acid that is hydrophobic; or
XX Thr; X7 = a residue capable of forming a bond (i) with a bridging amino
XX acid of X2, (ii) with X3 when X2 is absent, or (iii) with X4 when X2 and
XX X3 are absent, to conformationally restrain the peptide; X8 = NR3R4; OR5;
XX or 0-10 amino acids, optionally protected as an amide at the C-terminus;
XX R1, R3-R5 = H or alkyl; R2 = H, alkyl, alkyl-CO or phenyl-CO. The
XX peptides are useful for inhibiting bone resorption, angiogenesis or
XX restenosis, and for altering integrin receptor-mediated interactions,
XX especially alpha v beta 3 integrin receptor-inhibiting osteoclast binding
XX to a matrix. They may be used for reducing or inhibiting osteoclast binding
XX to a matrix. The present sequence represents an example of a circular
XX peptide disclosed in the specification. (Updated on 25-MAR-2003 to
XX correct PF field.)

XX Sequence 9 AA;
SQ

Query Match 87.5%; Score 28; DB 2; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.8e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 2 CRGDE 6

RESULT 60

AAV21792
ID AAV21792 standard; peptide; 9 AA.

XX AC AAV21792;

XX 10-SEP-1999 (first entry)
DT

XX Conserved peptide sequence of ICH-1 (caspase-2).
DE

XX Rev-caspase; cysteine protease; zymogen; caspase; cancer; gene therapy;
KW autoimmune disease; caspase-mediated apoptosis; neurodegenerative;
KW tumour cell; myocardial infarction; human.

XX Homo sapiens.
OS

XX WO9935277-A2.
PN

XX 15-JUL-1999.
PD

XX 11-JAN-1999; 99WO-US000632.
PF

XX 09-JAN-1998; 98US-0070987P.
PR

XX (UYJE-) UNIV JEFFERSON THOMAS.
PA

XX Alnemri ES;
PI

XX WPI; 1999-419353/35.
DR

XX New isolated nucleic acid molecule encoding a rev-caspase - used for
PT screening and identifying inhibitors or enhancers for treating cancer or
PT autoimmune disease.

XX Disclosure; Fig 6; 74pp; English.
PS

XX The invention relates to nucleic acid molecules encoding rev-caspases.
XX Rev-caspases are cysteine proteases that specifically cleave proteins
XX after Asp residues and is expressed as a zymogen, in which the small
XX subunit is N-terminal to a large subunit. A gene delivery vehicle
XX comprising a rev-caspase coding sequence is internalised by tumour cells.
XX cancer, where the gene delivery vehicle is used to treat autoimmune diseases.
XX The gene delivery vehicle can also be used to treat autoimmune diseases.
XX Cells transfected with a rev-caspase expressing vector can be used in
XX identification of inhibitors or enhancers of caspase-mediated apoptosis.
XX In vitro translated rev-caspase can be used to identify an inhibitor or
XX enhancer of caspase processing activity. Caspase inhibitors are useful
XX for treating neurodegenerative diseases as well as for inhibiting
XX apoptosis in the heart following myocardial infarction. Sequences
XX AAV21734 -AAV21795 represent conserved peptide sequences in various
XX caspases

XX Sequence 9 AA;
SQ

Query Match 87.5%; Score 28; DB 2; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.8e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 5 CRGDE 9

RESULT 61
AAW97010
ID AAW97010 standard; peptide; 9 AA.
XX
AC AAW97010;
XX
DT 26-APR-1999 (first entry)
XX
DE Peptide used to inhibit bone resorption, angiogenesis and restenosis.
XX
KW RGD-peptide; bone resorption; angiogenesis; restenosis;
KW integrin receptor mediated interaction; binding; osteoclast; bone.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT Modified-site 10 /note= "acetylated residue"
FT Modified-site 10 /note= "amidated residue"
XX
PN US5849865-A.
XX
PD 15-DEC-1998.
XX
PF 12-APR-1995; 95US-00421695.
XX
PR 13-APR-1994; 94US-00227316.
PR 08-SEP-1994; 94US-00303052.
XX
PA (LJOL-) LA JOLLA CANCER RES FOUND.
XX
PI Techopp JF, Mullen D, Ingram R, Cheng S;
XX
DR WPI; 1999-179486/15.
XX
PT RGD peptides - that modulate integrin-mediated cell-matrix binding.
XX
PS Claim 6; Col 138; 90pp; English.
XX
CC The present peptide represents a non-naturally occurring Arg-Gly-Asp
(RGD)-containing peptide. The peptide may be useful for reducing or
CC inhibiting bone resorption, angiogenesis or restenosis, and for altering
CC an integrin receptor mediated interaction. The peptide may also be used
CC to inhibit the binding of an osteoclast to bone
XX
SQ Sequence 9 AA;
Query Match 87.5%; Score 28; DB 2; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.8e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CRGDD 5
Db 2 CRGDE 6
RESULT 62
AAE09007
ID AAE09007 standard; peptide; 9 AA.
XX
AC AAE09007;
XX
DT 15-NOV-2001 (first entry)
XX
DE ICH-1 peptide fragment #3.
XX
KW Apoptic protease; mammalian ced-3 homologue 6; Mch6; cancer; ICH-1;
KW aspartate-specific cysteine protease; ASCP; apoptosis; therapy;
KW autoimmune disease; cerebellar degeneration; Alzheimer's disease;
KW cytostatic; Parkinson's disease; immunomodulator; antimicrobial;
KW viral infection; cell death-mediated disease; neuroprotective.
XX
XX

OS Unidentified.
XX
PN US6271361-B1.
XX
PD 07-AUG-2001.
XX
PF 25-FEB-1999; 99US-00257218.
XX
PR 29-MAY-1997; 97US-00865579.
XX
PA (UYJB-) UNIV JEFFERSON THOMAS.
XX
PI Alnemri ES, Fernandes-Alnemri T, Litwack G;
XX
DR WPI; 2001-528686/58.
XX
XX
PT New apoptotic genes and their apoptotic protease products, useful for
PT modulating apoptosis for the therapeutic treatment of human diseases,
PT e.g. cancers, autoimmune disease, Alzheimer's disease or Parkinson's
PT disease.
XX
PS Disclosure; Col 51; 36pp; English.
XX
CC The invention relates to an isolated gene encoding apoptic protease,
CC mammalian ced-3 homologue 6 (Mch6). Mch6 is a member of the aspartate-
CC specific cysteine protease (ASCP) family. Mch6 DNA and protein sequences
CC are useful for modulating apoptosis for the therapeutic treatment of
CC human diseases. Mch6 sequences are useful for upregulating apoptosis
CC (e.g. for treating cancers, autoimmune disease or viral infections) or
CC downregulating apoptosis (e.g. for treating Alzheimer's disease,
CC Parkinson's disease or cerebellar degeneration). The Mch6 sequence is
CC useful for diagnosing, treating or reducing the severity of cell death-
CC mediated diseases, as well as other diseases mediated by either increased
CC or decreased programmed cell death. The present amino acid sequence is a
CC fragment of ICH-1 peptide which belongs to NEDD-like subfamily
XX
SQ Sequence 9 AA;
Query Match 87.5%; Score 28; DB 4; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.8e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CRGDD 5
Db 5 CRGDE 9
RESULT 63
ABJ01296
ID ABJ01296 standard; peptide; 9 AA.
XX
AC ABJ01296;
XX
DT 18-SEP-2002 (first entry)
XX
DE Human caspase conserved region SEQ ID NO: 112.
XX
KW Human; caspase; rev-caspase; gene therapy; protease; apoptosis; cancer;
KW autoimmune disease; cytostatic; immunosuppressive.
XX
OS Homo sapiens.
XX
PN US6376226-B1.
XX
PD 23-APR-2002.
XX
PF 26-APR-2000; 2000US-00561756.
XX
PR 09-JAN-1998; 98US-0070897P.
PR 08-JAN-1999; 99US-00227721.
XX
PA (UYJB-) UNIV JEFFERSON THOMAS.
XX
XX

PI Alnemri ES;
XX
DR WPI; 2002-453146/48.
XX
PT New rev-caspases engineered to contain the small subunit fused in frame N
PT -terminal to the large subunit, which is in reverse order to the wild
PT type caspases, are useful to treat cancer and autoimmune diseases.
XX
XX Disclosure; Fig 6; 81pp; English.
XX
CC The present invention provides the protein and coding sequences of human
CC rev-caspase-3, uncleavable rev-caspase-3 and rev-caspase-6. The sequences
CC can be used in the gene therapy of cancer and autoimmune diseases. The
CC present sequence is a peptide described in the exemplification of the
CC invention
XX
SQ Sequence 9 AA;

Query Match 87.5%; Score 28; DB 5; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.8e+06; Indels 0; Gaps 0;
Matches 4; Conservative 1; Mismatches 0;

Qy 1 CRGDD 5
Db 5 CRGDE 9
|||||

RESULT 64
ADM96249
ID ADM96249 standard; peptide; 21 AA.
XX
AC ADM96249;
XX
DT 15-JUL-2004 (first entry)
XX
DE Human caspase-2 peptide related to apoptosis modulation.
XX
KW apoptotic cell death; apoptotically active site; human alpha-fetoprotein;
KW human serum albumin; immunosuppressive; virucide;
KW apoptosis signaling modulator; apoptotic regulatory pathway;
KW organ preservation; cell preservation; transplantation;
KW autoimmune disorder; immunodeficiency syndrome; viral infection;
KW cytotoxicity; chemotherapy; radiotherapy; neuronal cell apoptosis;
KW non-specific drug-induced apoptosis; oxidative stress-mediated apoptosis;
KW cultured cell apoptosis; caspase-2; human.
XX
OS Homo sapiens.
XX
XX WO2004033500-A1.
XX
PN 22-APR-2004.
XX
PD 07-OCT-2003; 2003WO-FI000735.
XX
PF 09-OCT-2002; 2002FI-00001798.
XX
PR (DUDI// DUDICH E I.
PA (SEME// SEMENKOVA L N.
PA (DUDI// DUDICH I V.
PA (TATU// TATULOV E B.
PA (ZUBO// ZUBOV D L.
PA (KORP// KORPELA T K.
XX
XX Dudich EI, Semenkova LN, Dudich IV, Tatulov EB, Zubov DL;
PI KorpeLA TK;
PI
XX WPI; 2004-340893/31.
DR
XX
XX New molecular structure of the recognition site of an anti-idiotypic
PT antibody against an apoptotically active site of human alpha-fetoprotein,
PT useful for preventing of autoimmune disorders and an immunodeficiency
PT syndrome.
PT
XX

PS Example 2; Page 13; 36pp; English.
XX
CC This invention relates to novel peptides capable of modulating apoptotic
CC cell death. In particular, the invention relates to the molecular
CC structure of the apoptotically active site of human alpha-fetoprotein and
CC human serum albumin. The invention may be useful for the development of
CC compounds with an immunosuppressive or virucide activity acting as
CC apoptosis signaling modulators. The peptide structures are useful for
CC suppressing apoptotic regulatory pathways in human and animal cells, for
CC increasing preservation of organs or cells within their transplantation,
CC for preventing of autoimmune disorders and an immunodeficiency syndrome,
CC induced by a viral infection, for lowering cytotoxic effects after chemo
CC or radiotherapy, for inhibiting neuronal cell apoptosis, non-specific
CC drug-induced apoptosis or oxidative stress-mediated apoptosis, and for
CC preventing apoptosis of cultured cells prepared for scientific or
CC technical purposes. The present sequence is that of a human caspase-2
CC peptide which was used in the exemplification of the invention.
XX
SQ Sequence 21 AA;

Query Match 87.5%; Score 28; DB 8; Length 21;
Best Local Similarity 80.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 12 CRGDE 16
|||||

RESULT 65
ABU61264
ID ABU61264 standard; peptide; 35 AA.
XX
AC ABU61264;
XX
DT 08-MAY-2003 (first entry)
XX
DE Human A domain from APOER2 #7.
XX
KW LDL-receptor class A domain; A domain; human; domain multimer;
KW multimer library; immuno-domain library.
XX
OS Homo sapiens.
XX
XX WO200288171-A2.
XX
PN 07-NOV-2002.
XX
PD 26-APR-2002; 2002WO-US013257.
XX
PF 26-APR-2001; 2001US-0286823P.
XX
PR 19-NOV-2001; 2001US-0337209P.
XX
PR 26-NOV-2001; 2001US-033359P.
XX
PR 18-APR-2002; 2002US-0374107P.
XX
PA (MAXY-) MAXYGEN INC.
XX
XX Kolkman JA, Stemmer WPC;
XX
XX WPI; 2003-111869/10.
XX
XX
PT Identifying a multimer that binds to a target molecule, comprises
PT identifying at least one monomer domain that bind to at least one target
PT molecule and linking the identified monomer domains to form a library of
PT multimers.
XX
XX Disclosure; Fig 10; 98pp; English.
XX
XX The invention relates to identifying a multimer that binds to a target
CC molecule, comprising identifying at least one monomer domain that binds
CC to at least one target molecule, linking the identified monomer domains
CC to form a library of multimers, each multimer comprising at least two
CC monomer domains, and screening the library of multimers for the ability

to bind to the first target molecule. Also included are: (1) a library of multimers formed by the method above (where each multimer comprises at least two monomer domains connected by a linker, and each monomer domain exhibits a binding specificity for a target molecule); (2) a polypeptide comprising: (a) the multimer selected from the novel method; or (b) at least two monomer domains separated by a heterologous linker, where each monomer domain specifically binds to a target molecule; (3) a polynucleotide encoding the multimer selected from the novel method; and (4) identifying hetero-immuno multimers that bind to a target molecule, comprising: (a) providing a library of immuno-domains; (b) screening the library of immuno-domains for affinity to a first target molecule; (c) providing a library of monomer domains; (d) screening the library of monomer domains for affinity to a first target molecule; (e) identifying at least one immuno-domain that binds to at least one target molecule; (f) identifying at least one monomer domain that binds to at least one target molecule; (g) linking the identified immuno-domain with the identified monomer domains to form a library of multimers, each multimer comprising at least two domains; (h) screening the library of multimers for the ability to bind to the first target molecule; and (i) identifying a multimer that binds to the first target molecule. The methods are useful for identifying multimers that bind to target molecules. The methods can also be used for selecting and optimising properties of discrete monomer domains and/or immuno-domains to create multimers. The multimers are useful for identifying the multimers with improved phenotype such as improved avidity or affinity or altered specificity for the target molecule. The polynucleotide, polypeptide and/or multimer are useful for preventing or treating a disease or disorder in a subject. The present sequence is a human LDL (low density lipoprotein) class A domain or simply an A domain used to design a library of A domain multimers of the invention

XX Sequence 35 AA;

Query Match 87.5%; Score 28; DB 6; Length 35;
Best Local Similarity 80.0%; Pred. No. 4.4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
| | | | |
Db 1 CRGDE 5

RESULT 66
ADP21486
ID ADP21486 standard; peptide; 35 AA.
XX
AC ADP21486;
XX

12-AUG-2004 (first entry)
XX Human LDL receptor A domain peptide SeqID 62.
XX monomer; multimer; combinatorial library; LDL receptor A domain;
KW immunosuppressive; human; immunodomain; low density lipoprotein.
XX Homo sapiens.

XX WO2004044011-A2.
XX 27-MAY-2004.
XX
XX 06-NOV-2003; 2003WO-US035664.
XX
XX 06-NOV-2002; 2002US-00289660.
XX
XX (AVID-) AVIDIA RES INST.
XX

XX Kolkman JA, Stemmer WPC;
XX WPI; 2004-411701/38.
XX
XX
XX
XX

Identifying multimers binding to targets, by screening a library of domains for affinity to the target, identifying domains binding to the

PT target, forming, screening a library of multimers and identifying multimer binding to the target.
XX
PS Disclosure; Fig 10; 161pp; English.

XX This invention relates to a novel method for identifying monomers and multimers that bind to a target molecule. Specifically, it refers to the generation of combinatorial libraries by screening a library of monomer and/or immunodomains to identify affinity for a target, and furthermore linking the identified monomer domains to a second monomer domain to form a library of multimers that still exhibit the ability to bind to the first target molecule. The present invention refers to monomer domains selected from the group consisting of EGF, EF, cadherin, C-type lectin, C2 and annexin domains, and in particular to a library of DNA sequences encoding the LDL receptor A domains. The generation of appropriate monomer/ multimer libraries is useful for creating antagonists to a target molecule, as well as for detection and other affinity based assays, such that the identified molecules can be used to increase the serum half-life of a molecule in the blood or for treating/ preventing a disease associated with the target molecule. In addition, they exhibit immunosuppressive activities and are useful for lowering the immune response in a mammal. This peptide sequence is a human low density lipoprotein (LDL) receptor class A-domain peptide of the invention.

XX Sequence 35 AA;

Query Match 87.5%; Score 28; DB 8; Length 35;
Best Local Similarity 80.0%; Pred. No. 4.4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
| | | | |
Db 1 CRGDE 5

RESULT 67
AAR13677
ID AAR13677 standard; protein; 39 AA.
XX
AC AAR13677;
XX

13-NOV-1991 (first entry)
XX Decorsin.
XX

Thrombin; factor Xa; cleavage site; receptor; GPIIb/IIIa; blood; clotting; infarction; stroke; platelet aggregation; leech.

XX Macrobodella decora.

XX Key Location/Qualifiers
FH Peptide 23..37
FT /label= 2 /note= "claim 3, page 54"
FT 28..37
FT /label= 1
FT /note= "claim 1, page 54"

XX WO9112270-A.

XX 22-AUG-1991.

XX 14-FEB-1990; 90US-00479829.

XX 14-FEB-1990; 90US-00479829.

XX 26-OCT-1990; 90US-00602847.

XX (GETH) GENENTECH INC.

XX Lazarus RA, Seymour JL;

XX WPI; 1991-267097/36.

XX N-PSDB; AAQ13325.

XX New platelet aggregation inhibitors - derived from species of leech and
 PT used to treat conditions where blood has increased tendency to clot.
 XX
 XX
 PS Disclosure; Fig 7; 71pp; English.
 XX
 CC Decorin has three disulfide bonds in its native (i.e. unreduced) form.
 CC Mol. wt. is less than 14.4 kD (SDS-PAGE) and 4384 D, calculated on the
 CC reduced protein. Decorin is about 27 times more potent than the
 CC pentapeptide GRGDV. A peptide at least 70%, pref. 80%, homologous to
 CC indicated peptide 1 and pref. indicated peptide 2 is functional as
 CC inhibitor of platelet aggregation. The prod. binds to the GPIIb/IIIa
 CC receptors so prevent binding of fibrinogen to platelets. It can be
 CC administered therapeutically (esp. to provide a concn. of 1nM-50 microM)
 CC to prevent abnormal blood clotting (e.g. in phlebitis, infarction,
 CC stroke, etc.), opt. together with a fibrinolytic agent). See also
 CC AAQ13324-26, AAR13676-78, AAR13672-74 and AAR13761-63
 XX
 SQ Sequence 39 AA;

Query Match 87.5%; Score 28; DB 2; Length 39;
 Best Local Similarity 80.0%; Pred. No. 4.9e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
 |:|||
 Db 7 CQGDD 11

RESULT 68
 AAR95440
 ID AAR95440 standard; protein; 39 AA.

AC AAR95440;
 XX
 XX 17-SEP-1996 (first entry)
 DT
 XX

DE Decorin platelet aggregation inhibitor.

XX Multifunctional protein; targeted hetero-association; protein assembly;
 KW decorin; platelet aggregation inhibitor; antiaggregant; antithrombotic;
 KW leech.
 XX

OS Macrobella decora.

XX WO9613583-A2.

PN 09-MAY-1996.

PD 20-OCT-1995; 95WO-EP004117.

PF 20-OCT-1994; 94EP-00116558.

PR (MORP-) MORPHOSYS GES PROTEINOPTIMIERUNG MBH.

PA Pack P, Lupas A;

PI WPI; 1996-239496/24.

DR Targetted hetero-association of recombinant proteins to multi-functional
 XX complexes - useful for therapeutic and diagnostic purposes.
 PT

XX Example 7; Fig 23; 34pp; English.

XX Decorin (AAR95440), from the leech Macrobella decora, acts as a potent
 CC antagonist of the platelet glycoprotein IIb-IIIa. 2 Gene cassettes
 CC (AAT15273 and AAT15274) respectively coding for the N-terminal (AAR95441)
 CC and C-terminal (AAR95442) portion of decorin can be fused to an
 CC association segment used in a novel method for the targeted hetero-
 CC association of recombinant proteins to multi-functional complexes. A
 CC multivalent decorin complex combined with an anti-fibrin antibody
 CC fragment can act as a powerful antithrombotic agent
 XX

SQ Sequence 39 AA;

Query Match 87.5%; Score 28; DB 2; Length 39;
 Best Local Similarity 80.0%; Pred. No. 4.9e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
 |:|||
 Db 7 CQGDD 11

RESULT 69

AAB92369
 ID AAB92369 standard; peptide; 39 AA.

AC AAB92369;

XX 22-JUN-2001 (first entry)
 DT
 XX

DE Miscellaneous peptide SEQ ID NO:1545.

XX Protection; endogenous therapeutic peptide; peptidase; conjugation;
 KW blood component; modification; succinimidyl; maleimido group; amino;
 KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.
 XX

OS Homo sapiens.

OS Synthetic.

XX WO200069900-A2.

PN 23-NOV-2000.

PD 17-MAY-2000; 2000WO-US013576.

PF 17-MAY-1999; 99US-0134406P.

PR 10-SEP-1999; 99US-0153406P.

PR 15-OCT-1999; 99US-0159783P.

XX (CONJ-) CONJUCHEM INC.

XX Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;

XX WPI; 2001-112059/12.

XX Modifying and attaching therapeutic peptides to albumin prevents
 PT peptidase degradation, useful for increasing length of in vivo activity.

XX Disclosure; Page 710; 733pp; English.

XX The present invention describes a modified therapeutic peptide (I)
 CC comprising a therapeutically active amino acid region (III) and a
 CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to
 CC a less therapeutically active amino acid region (IV), which covalently
 CC bonds with amino/hydroxyl/thiol groups on blood components to form a
 CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
 CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
 CC factors and neurotransmitters, to protect them from peptidase activity in
 CC vivo for the treatment of various disorders. Endogenous therapeutic
 CC peptides are not suitable as drug candidates as they require frequent
 CC administration due to rapid degradation by peptidases in the body.
 CC Modifying and attaching therapeutic peptides to albumin prevents or
 CC reduces the action of peptidases to increase length of activity (half
 CC life) and specificity as bonding to large molecules decreases
 CC intracellular uptake and interference with physiological processes.
 CC AAB90829 to AAB92441 represent peptides which can be used in the
 CC exemplification of the present invention
 XX

SQ Sequence 39 AA;

Query Match 87.5%; Score 28; DB 4; Length 39;
 Best Local Similarity 80.0%; Pred. No. 4.9e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
|:|
Db 7 CQGGD 11

RESULT 70
AAW20043
ID AAR95442 standard; protein; 41 AA.
AC AAR95442;
XX 17-SEP-1996 (first entry)
DT
XX Decorsin C-terminal portion.
DE
XX Multifunctional protein; targetted hetero-association; protein assembly;
KW decorsin; platelet aggregation inhibitor; antiaggregant; antithrombotic;
KW leech; Macrobodella decora.
XX
XX Synthetic.
OS
XX WO9613583-A2.
PN
XX
XX 09-MAY-1996.
PD
XX
XX 20-OCT-1995; 95WO-EP004117.
PF
XX
XX 20-OCT-1994; 94EP-00116558.
PR
XX
XX (MORP-) MORPHOSYS GES PROTEINOPTIMIERUNG MBH.
PA
XX
XX Pack P, Lupas A;
PI
XX
XX MPI; 1996-239496/24.
DR
XX N-PSDB; AAT15274.
DR
XX
XX Targetted hetero-association of recombinant proteins to multi-functional
PT complexes - useful for therapeutic and diagnostic purposes.
PT
XX
XX Example 7; Fig 25; 34pp; English.
PS
XX
XX Decorsin (AAR95440), from the leech Macrobodella decora, acts as a potent
CC antagonist of the platelet glycoprotein IIB-IIIa. 2 Gene cassettes
CC (AAT15273 and AAT15274) respectively coding for the N-terminal (AAR95441)
CC and C-terminal (AAR95442) portion of decorsin can be fused to an
CC association segment used in a novel method for the targetted hetero-
CC association of recombinant proteins to multi-functional complexes. A
CC multivalent decorsin complex combined with an anti-fibrin antibody
CC fragment can act as a powerful antithrombotic agent
XX
XX Sequence 41 AA;

Query Match 87.5%; Score 28; DB 2; Length 41;
Best Local Similarity 80.0%; Pred. No. 5.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
|:|
Db 9 CQGGD 13

RESULT 71
AAW20043
ID AAW20043 standard; protein; 41 AA.
XX
AC AAW20043;
XX
XX 04-SEP-1997 (first entry)
DT
XX Decorsin platelet aggregation inhibitor with flanking linker residues.
DE
XX Multimerisation; self assembly; functional domain; linker; folding;
KW multimerisation domain; post-translational modification; secretion;
DB

KW interleukin-2; TAFII131; TAFII80; TATA box binding associated factor; p53;
KW histone; H3; H4; thrombospondin; TSP-4; platelet factor; PF4;
KW cartilage oligomeric protein; COMP.
XX
OS Synthetic.
XX
XX Key Location/Qualifiers
FT Misc-difference 3. .41
FT
FT
XX
XX WO9637621-A2.
PN
XX
XX 28-NOV-1996.
PD
XX
XX 23-MAY-1996; 96WO-EP002230.
PF
XX
XX 23-MAY-1995; 95EP-00107914.
PR
XX
XX (MORP-) MORPHOSYS GES PROTEINOPTIMIERUNG MBH.
PA
XX
XX Pack P, Hoess A;
PI
XX
XX MPI; 1997-021226/02.
DR
XX N-PSDB; AAT71291.
DR
XX
XX Multimerisation devices for self assembly of multifunctional proteins -
PT used to express recombinant multivalent poly:peptide(s) by incorporation
PT in a cistron encoding the protein.
PT
XX
XX Example 5; Fig 15; 64pp; English.
PS
XX
XX AAW20043 shows human decorsin, a platelet aggregation inhibitor, flanked
CC by residues of surrounding linker peptides. The sequence is the product
CC of an expression cassette that was incorporated into a larger DNA
CC cassette as a C-terminal functional domain. The larger cassette
CC comprises, in 5' to 3' direction, a 1st functional domain; a 1st linker
CC sequence; a multimerisation device; a 2nd linker sequence; and a 2nd
CC functional domain. The multimerisation device allows the combination of
CC two or more functional domains in a structure which is capable of self-
CC multimerisation (at least trimerisation). Functional domains may, for
CC example, bind to a defined target, catalyse a reaction, block a receptor
CC binding site, inhibit the action of another protein or bind to a metal
CC ion. Multimerisation domains from p53, platelet factor 4, thrombospondin,
CC TSP-4, TATA box binding associated factors and cartilage oligomeric
CC protein may be used. The multifunctional proteins can be prepared using
CC standard recombinant micro-organisms, even though the molecular weight of
CC the assembled protein exceeds that of the proteins commonly expressed in
CC bacteria. They have low immunogenicity in humans and carry two or more
CC functions in a single multimeric structure. Use of a combination of in
CC vivo expression and in vitro synthesis overcomes prior art problems due
CC to the differences in folding, secretion and post-translational
CC modifications for different polypeptides in different hosts
XX
XX Sequence 41 AA;

Query Match 87.5%; Score 28; DB 2; Length 41;
Best Local Similarity 80.0%; Pred. No. 5.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
|:|
Db 9 CQGGD 13

RESULT 72
AAR95441
ID AAR95441 standard; protein; 42 AA.
XX
XX AAR95441;
AC
XX
XX 17-SEP-1996 (first entry)
DT
XX Decorsin N-terminal portion.
DE

XX Multifunctional protein; targeted hetero-association; protein assembly;
 KW decorsin; platelet aggregation inhibitor; antiaggregant; antithrombotic;
 KW leech; Macrobodella decora.
 XX Synthetic.
 OS
 XX WO9613583-A2.
 PN
 XX 09-MAY-1996.
 PD
 XX 20-OCT-1995; 95WO-EP004117.
 XX
 XX 20-OCT-1994; 94EP-00116558.
 XX
 XX (MORP-) MORPHOSYS GES PROTEINOPTIMIERUNG MBH.
 PA
 XX Pack P, Lupas A;
 PI
 XX WPI; 1996-239496/24.
 DR
 XX N-PSDB; AAT15273.
 DR
 XX Targetted hetero-association of recombinant proteins to multi-functional
 PT complexes - useful for therapeutic and diagnostic purposes.
 PT
 XX Example 7; Fig 24; 34pp; English.
 PS
 XX Decorsin (AAR95440), from the leech Macrobodella decora, acts as a potent
 CC antagonist of the platelet glycoprotein IIb-IIIa. 2 Gene cassettes
 CC (AAT15273 and AAT15274) respectively coding for the N-terminal (AAR95441)
 CC and C-terminal (AAR95442) portion of decorsin can be fused to an
 CC association segment used in a novel method for the targetted hetero-
 CC association of recombinant proteins to multi-functional complexes. A
 CC multivalent decorsin complex combined with an anti-fibrin antibody
 CC fragment can act as a powerful antithrombotic agent
 XX
 SQ Sequence 42 AA;
 Query Match 87.5%; Score 28; DB 2; Length 42;
 Best Local Similarity 80.0%; Pred. No. 5.2e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 CRGDD 5
 Db |::|||
 9 CQGDD 13
 RESULT 73
 AAW20042
 ID AAW20042 standard; protein; 42 AA.
 AC
 XX AAW20042;
 XX
 DT 04-SEP-1997 (first entry)
 XX
 DE Decorsin platelet aggregation inhibitor with flanking linker residues.
 XX
 XX Multimerisation; self assembly; functional domain; linker; folding;
 KW multimerisation domain; post-translational modification; secretion;
 KW interleukin-2; TAFI131; TAFI180; TATA box binding associated factor; p53;
 KW histone; H3; H4; thrombospondin; TSP-4; platelet factor; PF4;
 KW cartilage oligomeric protein; COMP; ss.
 XX
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 3..41
 FT /note= "decorsin"
 FT
 XX WO9637621-A2.
 PN
 XX 28-NOV-1996.
 PD
 XX

PF 23-MAY-1996; 96WO-EP002230.
 XX
 PR 23-MAY-1995; 95EP-00107914.
 XX
 PA (MORP-) MORPHOSYS GES PROTEINOPTIMIERUNG MBH.
 XX
 XX Pack P, Hoess A;
 PI
 XX WPI; 1997-021236/02.
 DR
 XX N-PSDB; AAT71290.
 DR
 XX Multimerisation devices for self assembly of multifunctional proteins -
 PT used to express recombinant multivalent poly.peptide(s) by incorporation
 PT in a cistron encoding the protein.
 XX
 XX Example 5; Fig 14; 64pp; English.
 PS
 XX AAW20042 shows human decorsin, a platelet aggregation inhibitor, flanked
 CC by residues of surrounding linker peptides. The sequence is the product
 CC of an expression cassette that was incorporated into a larger DNA
 CC cassette as an N-terminal functional domain. The larger cassette
 CC comprises, in 5' to 3' direction, a 1st functional domain; a 1st linker
 CC sequence; a multimerisation device; a 2nd linker sequence; and a 2nd
 CC functional domain. The multimerisation device allows the combination of
 CC two or more functional domains in a structure which is capable of self-
 CC multimerisation (at least trimerisation). Functional domains may, for
 CC example, bind to a defined target, catalyse a reaction, block a receptor
 CC binding site, inhibit the action of another protein or bind to a metal
 CC ion. Multimerisation domains from p53, platelet factor 4, thrombospondin,
 CC TSP-4, TATA box binding associated factors and cartilage oligomeric
 CC protein may be used. The multifunctional proteins can be prepared using
 CC standard recombinant micro-organisms, even though the molecular weight of
 CC the assembled protein exceeds that of the proteins commonly expressed in
 CC bacteria. They have low immunogenicity in humans and carry two or more
 CC functions in a single multimeric structure. Use of a combination of in
 CC vivo expression and in vitro synthesis overcomes prior art problems due
 CC to the differences in folding, secretion and post-translational
 CC modifications for different polypeptides in different hosts
 XX
 SQ Sequence 42 AA;
 Query Match 87.5%; Score 28; DB 2; Length 42;
 Best Local Similarity 80.0%; Pred. No. 5.2e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 CRGDD 5
 Db |::|||
 9 CQGDD 13
 RESULT 74
 AAR13676
 ID AAR13676 standard; protein; 45 AA.
 AC
 XX AAR13676;
 XX
 DT 13-NOV-1991 (first entry)
 XX
 DE Decorsin fusion protein.
 XX
 KW Thrombin; factor Xa; cleavage site; receptor; GRP1BIIa; blood; clotting;
 KW infarction; stroke; platelet aggregation; leech.
 XX
 OS Macrobodella decora.
 XX
 XX Key Location/Qualifiers
 FH Cleavage-site 1..6
 FT /note= "cleavable by thrombin and factor Xa"
 FT Protein 7..45
 FT /label= decorsin
 FT Peptide 29..43
 FT /label= 2
 FT /note= "claim 3, page 54"
 FT

FT Peptide 34. 43
FT /label= 1
FT /note= "claim 1, page 54"
XX
PN WO9112270-A.
XX
PD 22-AUG-1991.
XX
XX 14-FEB-1990; 90US-00479829.
XX
XX 14-FEB-1990; 90US-00479829.
PR 26-OCT-1990; 90US-00602847.
XX
XX (GETH) GENENTECH INC.
PA
XX Lazarus RA, Seymour JL;
XX WPI; 1991-267097/36.
DR N-PSDB; AAQ13324.
XX
XX New platelet aggregation inhibitors - derived from species of leech and
PT used to treat conditions where blood has increased tendency to clot.
XX
XX Disclosure; Fig 6; 71pp; English.
XX
XX The sequence comprises decorin and a 6 amino acid cleavage site amenable
CC to cleavage by thrombin and factor Xa. This cleavage site is used to
CC generate decorin from a fusion protein (constructed e.g. for secretion
CC into the periplasmic space or for intracellular expression). A peptide at
CC least 70%, pref. 80%, homologous to indicated peptide 1 and pref.
CC indicated peptide 2 is functional as inhibitor of platelet aggregation.
CC The prod. binds to the GPIIb/IIIa receptors so prevent binding of
CC fibrinogen to platelets. It can be administered therapeutically (esp. to
CC provide a concn. of 1m-50 microm) to prevent abnormal blood clotting
CC (e.g. in phlebitis, infarction, stroke, etc.), opt. together with a
CC fibrinolytic agent). See also AAQ13324-26, AAR13676-78, AAR13672-74 and
XX AAR13761-63
XX
SQ Sequence 45 AA;

Query Match 87.5%; Score 28; DB 2; Length 45;
Best Local Similarity 80.0%; Pred. No. 5.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 CRGDD 5
DB 13 CQGDD 17

RESULT 75
AAE08949
ID AAE08949 standard; peptide; 46 AA.
XX
AC AAE08949;
XX

DT 15-NOV-2001 (first entry)
XX

DE ICH-1 peptide of NEDD-like ASCP subfamily.
XX

KW Apoptotic protease; mammalian ced-3 homologue 6; Mch6; cancer; ICH-1;
KW aspartate-specific cysteine protease; ASCP; apoptosis; therapy;
KW autoimmune disease; cerebellar degeneration; Alzheimer's disease;
KW cytostatic; Parkinson's disease; immunomodulator; antimicrobial;
KW viral infection; cell death-mediated disease; neuroprotective.
XX

OS Unidentified.
XX

XX US6271361-B1.
PN

XX 07-AUG-2001.
PD

XX 25-FEB-1999; 99US-00257218.
PF

XX

PR 29-MAY-1997; 97US-00865579.
XX
XX (UYJE-) UNIV JEFFERSON THOMAS.
PA
XX Alnemri ES, Fernandes-Alnemri T, Litwack G;
PI WPI; 2001-528686/58.
XX
XX New apoptotic genes and their apoptotic protease products, useful for
PT modulating apoptosis for the therapeutic treatment of human diseases,
PT e.g. cancers, autoimmune disease, Alzheimer's disease or Parkinson's
PT disease.
XX
XX Example 1; Fig 2; 36pp; English.
PS
XX The invention relates to an isolated gene encoding apoptotic protease,
CC mammalian ced-3 homologue 6 (Mch6). Mch6 is a member of the aspartate-
CC specific cysteine protease (ASCP) family. Mch6 DNA and protein sequences
CC are useful for modulating apoptosis for the therapeutic treatment of
CC human diseases. Mch6 sequences are useful for upregulating apoptosis
CC (e.g. for treating cancers, autoimmune disease or viral infections) or
CC downregulating apoptosis (e.g. for treating Alzheimer's disease,
CC Parkinson's disease or cerebellar degeneration). The Mch6 sequence is
CC useful for diagnosing, treating or reducing the severity of cell death-
CC mediated diseases, as well as other diseases mediated by either increased
CC or decreased programmed cell death. The present amino acid sequence is
CC ICH-1 peptide which belongs to NEDD-like ASCP subfamily
XX
SQ Sequence 46 AA;
Query Match 87.5%; Score 28; DB 4; Length 46;
Best Local Similarity 80.0%; Pred. No. 5.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 CRGDD 5
DB 17 CRGDE 21

Search completed: September 7, 2005, 20:09:43
Job time : 196 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2005, 19:58:16 ; Search time 39 Seconds
(without alignments)
12.335 Million cell updates/sec

Title: US-10-812-238B-41

Perfect score: 32

Sequence: 1 CRGDD 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

1: Pir1.*

2: Pir2.*

3: Pir3.*

4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	100.0	255	2	G64320
2	32	100.0	289	2	T44599
3	32	100.0	357	2	AG1072
4	32	100.0	363	2	G95937
5	32	100.0	400	1	A39130
6	32	100.0	400	1	A00756
7	32	100.0	400	2	AB0617
8	32	100.0	400	2	F90755
9	32	100.0	400	2	D85619
10	32	100.0	430	2	S75313
11	32	100.0	1116	2	T38073
12	29	90.6	264	2	T47183
13	29	90.6	296	2	A95408
14	29	90.6	413	2	F96743
15	29	90.6	732	2	T16422
16	29	90.6	761	2	F82078
17	28	87.5	39	2	A36453
18	28	87.5	66	1	DNVPS
19	28	87.5	74	1	DNVPS
20	28	87.5	182	2	I67436
21	28	87.5	234	2	S76342
22	28	87.5	416	2	B55649
23	28	87.5	423	2	B82611
24	28	87.5	427	2	T14237
25	28	87.5	431	2	S77330
26	28	87.5	435	2	A54821
27	28	87.5	452	2	JC6507
28	28	87.5	470	2	E72376
29	28	87.5	490	2	T24497

30	28	87.5	521	2	T18896
31	28	87.5	548	2	T22137
32	28	87.5	572	2	T51525
33	28	87.5	632	2	E81684
34	28	87.5	635	2	T00602
35	28	87.5	636	2	D84823
36	28	87.5	640	2	F71527
37	28	87.5	657	2	T00859
38	28	87.5	664	2	JH0770
39	28	87.5	730	2	E70798
40	28	87.5	842	2	C83177
41	28	87.5	1267	1	MXR31
42	28	87.5	1267	1	MXR32
43	28	87.5	1267	1	MXR33
44	28	87.5	1323	2	T37533
45	28	87.5	1806	2	T23298
46	28	87.5	2043	2	T18524
47	28	87.5	2098	2	T18397
48	28	87.5	3036	2	T18995
49	27	84.4	57	1	T1FHP
50	27	84.4	130	2	A13045
51	27	84.4	138	2	E95330
52	27	84.4	147	2	C95969
53	27	84.4	153	2	A60585
54	27	84.4	164	2	H64803
55	27	84.4	219	2	A70734
56	27	84.4	220	2	A44805
57	27	84.4	224	2	F64859
58	27	84.4	255	2	T15553
59	27	84.4	263	2	C56084
60	27	84.4	276	2	AF0402
61	27	84.4	299	2	F85835
62	27	84.4	299	2	D90990
63	27	84.4	299	2	E64975
64	27	84.4	323	2	G83892
65	27	84.4	326	1	GOVZML
66	27	84.4	333	2	B42476
67	27	84.4	333	2	S15309
68	27	84.4	333	2	AC0766
69	27	84.4	343	2	H70517
70	27	84.4	346	2	H70722
71	27	84.4	370	2	F84550
72	27	84.4	399	2	A84972
73	27	84.4	504	2	C64398
74	27	84.4	507	2	AD3382
75	27	84.4	553	2	E27793
76	27	84.4	602	1	S46676
77	27	84.4	603	2	JC6128
78	27	84.4	603	2	JC1282
79	27	84.4	605	2	JC5239
80	27	84.4	605	2	A41915
81	27	84.4	671	2	JC2029
82	27	84.4	684	2	T48357
83	27	84.4	721	2	H82528
84	27	84.4	914	2	B96592
85	27	84.4	940	2	T01854
86	27	84.4	986	2	T25512
87	27	84.4	1041	2	PQ0442
88	27	84.4	1058	2	A48195
89	27	84.4	1202	2	PQ0440
90	27	84.4	1283	2	T13799
91	27	84.4	1470	2	S45323
92	27	84.4	1475	2	T29809
93	27	84.4	1820	2	S71853
94	27	84.4	1847	2	T30558
95	27	84.4	2042	2	T18399
96	27	84.4	2212	2	T28157
97	27	84.4	2410	1	JQ1948
98	27	84.4	2412	1	JQ1537
99	27	84.4	2647	2	T28161
100	26	81.2	45	2	T05477
101	26	81.2	57	2	B38180
102	26	81.2	64	1	WMIH82

glypican 1 precurs
hypothetical prote
1-deoxyxylulose-5-
hypothetical prote
hypothetical prote
probable transket
hypothetical prote
probable atp-depen
fasciclin IV precu
probable phosphotr
lambda 3 protein -
lambda 3 protein -
coronin-like prote
hypothetical prote
protein CTRP - mal
scavenger receptor
hypothetical prote
proteinase inhibit
hypothetical prote
HmrR heavy metal d
probable transcrip
thyroglobulin - sl
ybfP protein - Bsc
probable cutinase
eggshell protein p
Putative lambdoid
hypothetical prote
interleukin-lbeta
transketolase (EC
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
T2 protein - myxom
hypothetical prote
hypothetical prote
probable glycoasyl
probable lppd prot
probable transfera
hypothetical prote
nicotinate phospho
hypothetical prote
ABC transporter-as
methyl coenzyme M
probable phosphoes
insulin-like growt
insulin-like growt
insulin-like growt
insulin-like growt
annexin - chicken
hypothetical prote
L-ascorbate oxidas
hypothetical prote
hypothetical prote
hypothetical prote
polyprotein - barl
ubiquitin-protein
polyprotein - barl
neurxin IV - frui
genome polyprotein
hypothetical prote
genome polyprotein
resistance protein
variant-specific s
erythrocyte membra
genome polyprotein
genome polyprotein
hypothetical prote
hypothetical prote
3b protein - avian

103	26	81.2	64	1	WMHBS	3b protein - avian	176	26	81.2	295	2	S05006	calcium channel al
104	26	81.2	64	1	WMHBS	3b protein - avian	177	26	81.2	302	2	A96661	unknown protein, 8
105	26	81.2	87	2	Q01683	U49A protein - hu	178	26	81.2	303	2	B84744	hypothetical prote
106	26	81.2	88	2	G84393	hypothetical prote	179	26	81.2	304	2	T42554	tegument protein -
107	26	81.2	102	2	A72518	hypothetical prote	180	26	81.2	308	2	D89950	porphobilinogen de
108	26	81.2	103	2	S62052	hypothetical prote	181	26	81.2	309	1	KIECRB	ribokinase (EC 2.7
109	26	81.2	104	2	D42446	hypothetical prote	182	26	81.2	309	2	AB0952	ribokinase (import
110	26	81.2	125	2	D97478	lipA protein (mpo	183	26	81.2	309	2	F91215	ribokinase (import
111	26	81.2	127	2	S01286	wheny acidic protei	184	26	81.2	309	2	G86061	ribokinase (import
112	26	81.2	137	2	T33210	hypothetical prote	185	26	81.2	311	2	F84030	porphobilinogen de
113	26	81.2	139	2	S58234	lipA protein - Rhi	186	26	81.2	311	2	B56084	interleukin-1beta
114	26	81.2	142	2	I41313	hypothetical prote	187	26	81.2	312	2	A61183	hypothetical prote
115	26	81.2	142	2	AD2696	lipA protein (mpo	188	26	81.2	313	2	T18092	hypothetical prote
116	26	81.2	145	2	F82133	hypothetical prote	189	26	81.2	315	2	T10613	hypothetical prote
117	26	81.2	149	2	S12913	cystatin - fruit f	190	26	81.2	321	2	H87683	transcription regu
118	26	81.2	152	2	I41310	hypothetical prote	191	26	81.2	323	2	T27640	hypothetical prote
119	26	81.2	152	2	D89753	protein Fl1C7.2 [i	192	26	81.2	325	2	C95952	hypothetical prote
120	26	81.2	156	2	B23861	tissue kallikrein	193	26	81.2	326	2	A97751	hypothetical prote
121	26	81.2	158	2	T04875	hypothetical prote	194	26	81.2	336	2	AH2040	porphobilinogen de
122	26	81.2	161	2	AC0893	conserved hypotet	195	26	81.2	336	2	E69845	thiamin biosynthes
123	26	81.2	163	2	S43240	troponin C - commo	196	26	81.2	337	2	T49627	probable DNA-direc
124	26	81.2	164	2	JM0061	troponin C - Flori	197	26	81.2	338	2	H83228	hypothetical prote
125	26	81.2	164	2	JW0060	troponin C - commo	198	26	81.2	342	2	T22167	hypothetical prote
126	26	81.2	168	2	T75542	hypothetical prote	199	26	81.2	349	2	I36944	haptoglobin - chim
127	26	81.2	173	2	T19469	hypothetical prote	200	26	81.2	350	2	AD3560	methylated-DNA-lpr
128	26	81.2	174	2	T15658	hypothetical prote	201	26	81.2	350	2	T25451	transforming growt
129	26	81.2	176	2	E96754	Similar to downy m	202	26	81.2	352	2	C96643	hypothetical prote
130	26	81.2	182	2	T20533	hypothetical prote	203	26	81.2	356	2	AH2431	oxidoreductase all
131	26	81.2	182	2	H69026	ribosomal protein	204	26	81.2	360	2	F69342	DNA primase homolo
132	26	81.2	187	2	B82857	pre-pilin like lea	205	26	81.2	360	2	A56066	basic helix-loop-h
133	26	81.2	188	2	B32340	tissue kallikrein	206	26	81.2	362	2	H69958	aminomethyltransfe
134	26	81.2	188	2	G86226	hypothetical prote	207	26	81.2	367	2	JE0104	testicular serine
135	26	81.2	193	2	E95340	hypothetical prote	208	26	81.2	371	2	F86546	hypothetical prote
136	26	81.2	196	2	T29016	hypothetical prote	209	26	81.2	371	2	C72077	conserved hypotet
137	26	81.2	205	2	T37940	hypothetical prote	210	26	81.2	371	2	D84186	hypothetical prote
138	26	81.2	208	2	T04796	hypothetical prote	211	26	81.2	373	2	PW0042	activin - fruit fl
139	26	81.2	209	2	H86259	protein T12C24.14	212	26	81.2	374	1	S35669	alcohol dehydrogen
140	26	81.2	213	2	S43723	lactase (EC 3.2.1.	213	26	81.2	374	1	DEHOAS	alcohol dehydrogen
141	26	81.2	213	2	S43722	lactase (EC 3.2.1.	214	26	81.2	374	1	S48157	alcohol dehydrogen
142	26	81.2	214	2	S65052	plastil-specific pr	215	26	81.2	375	1	A35837	alcohol dehydrogen
143	26	81.2	219	2	A53305	pentose-5-phosphat	216	26	81.2	375	1	A38405	alcohol dehydrogen
144	26	81.2	220	2	S70766	cathepsin D inhibi	217	26	81.2	375	1	A48257	alcohol dehydrogen
145	26	81.2	221	2	E96591	hypothetical prote	218	26	81.2	375	1	A49107	alcohol dehydrogen
146	26	81.2	222	2	S26994	cellulose 1,4-beta	219	26	81.2	375	1	DECHA1	alcohol dehydrogen
147	26	81.2	223	2	C69215	conserved hypotet	220	26	81.2	375	1	DEHOAL	alcohol dehydrogen
148	26	81.2	224	2	E90908	probable tail asse	221	26	81.2	375	1	DEMSAA	alcohol dehydrogen
149	26	81.2	226	2	H84237	hypothetical prote	222	26	81.2	375	1	I60970	alcohol dehydrogen
150	26	81.2	227	2	S77594	hypothetical prote	223	26	81.2	375	1	I60973	alcohol dehydrogen
151	26	81.2	228	2	S26995	cellulose 1,4-beta	224	26	81.2	375	1	S29343	alcohol dehydrogen
152	26	81.2	228	2	G70902	probable lprH prot	225	26	81.2	375	1	S62638	alcohol dehydrogen
153	26	81.2	228	2	T47425	NAC domain-like pr	226	26	81.2	375	1	S62639	alcohol dehydrogen
154	26	81.2	236	2	S26993	cellulose 1,4-beta	227	26	81.2	375	1	S66272	alcohol dehydrogen
155	26	81.2	240	2	A46587	extracellular hemo	228	26	81.2	376	1	A26468	occlusion-derived
156	26	81.2	243	2	B87631	enoyl-CoA hydratase	229	26	81.2	376	2	F72868	hypothetical prote
157	26	81.2	255	2	A84723	hypothetical prote	230	26	81.2	376	2	AD2697	3beta-hydroxy-Belt
158	26	81.2	258	2	C70968	probable ribD - My	231	26	81.2	377	2	I50147	Wnt inhibitory fac
159	26	81.2	258	2	T45931	hypothetical prote	232	26	81.2	378	2	B59180	phosphonate metabo
160	26	81.2	259	2	WMMS28	complement factor	233	26	81.2	378	2	AB2083	Similar to part of
161	26	81.2	259	2	T01008	ribonuclease (EC 3	234	26	81.2	380	2	C96754	hypothetical prote
162	26	81.2	259	2	E84427	hypothetical prote	235	26	81.2	380	2	T32163	hypothetical prote
163	26	81.2	262	1	JC4803	venombin A (EC 3.4	236	26	81.2	382	2	G85234	hypothetical prote
164	26	81.2	263	1	I55608	complement factor	237	26	81.2	383	2	A56084	interleukin-1beta
165	26	81.2	265	2	T04834	hypothetical prote	238	26	81.2	384	2	E75587	probable oxidoredu
166	26	81.2	266	2	T10609	hypothetical prote	239	26	81.2	384	2	A11843	hypothetical prote
167	26	81.2	268	2	T31699	hypothetical prote	240	26	81.2	389	2	T04006	11beta-hydroxyster
168	26	81.2	268	2	F87076	probable reductase	241	26	81.2	389	2	T27085	hypothetical prote
169	26	81.2	270	2	C96731	unknown protein F5	242	26	81.2	390	2	AD1035	probable major tai
170	26	81.2	270	2	F86177	protein Fl9P19.1 [243	26	81.2	390	2	AG0929	probable major tai
171	26	81.2	273	2	T10610	hypothetical prote	244	26	81.2	390	2	AI0836	probable bacteriop
172	26	81.2	274	2	T31943	hypothetical prote	245	26	81.2	392	1	A30100	serine proteinase
173	26	81.2	284	2	G96660	protein F2K11.8 [i	246	26	81.2	401	2	E88613	protein T2/29.9 [i
174	26	81.2	285	2	S72787	hypothetical prote	247	26	81.2	404	2	A42677	interleukin-1 beta
175	26	81.2	288	2	D83184	conserved hypotet	248	26	81.2	407	2	C97234	probable glycosyl

249	26	81.2	408	2	T25376	hypothetical prote	322	26	81.2	661	1	KPHU13	coagulation factor
250	26	81.2	408	2	G97194	ornithine acetyltr	323	26	81.2	663	2	G81384	succinate dehydrog
251	26	81.2	410	2	AF1974	hypothetical prote	324	26	81.2	665	2	T06082	protein kinase hom
252	26	81.2	417	2	B69254	3-ketoacyl-CoA thi	325	26	81.2	675	2	D85065	receptor protein k
253	26	81.2	423	2	T14526	S-locus-specific g	326	26	81.2	676	2	T47526	protein kinase-lik
254	26	81.2	423	2	AH0053	conserved hypothet	327	26	81.2	678	2	T04832	probable serine/th
255	26	81.2	424	2	F84729	hypothetical prote	328	26	81.2	688	1	A59271	Ra-reactive factor
256	26	81.2	426	2	T14527	S-locus-specific g	329	26	81.2	686	2	T05353	hypothetical prote
257	26	81.2	428	2	S35215	sodium channel pro	330	26	81.2	689	1	JW0107	very-long-chain ac
258	26	81.2	429	2	S48172	variable surface g	331	26	81.2	703	2	H86349	protein F8K7.9 (im
259	26	81.2	430	2	E86188	hypothetical prote	332	26	81.2	705	2	A70669	probable acid-CoA
260	26	81.2	430	2	B84212	hypothetical prote	333	26	81.2	705	2	A86350	F8K7.10 protein -
261	26	81.2	433	2	T14392	S-locus glycoprote	334	26	81.2	707	2	T35196	probable DNA gyras
262	26	81.2	441	2	T49819	hypothetical prote	335	26	81.2	707	2	F86925	probable acyl-CoA
263	26	81.2	447	2	C82958	probable two-compo	336	26	81.2	708	1	TFBOL	lactotransferrin p
264	26	81.2	450	2	A34169	alpha-2A-adrenergic	337	26	81.2	708	2	JC2323	lactoferrin - goat
265	26	81.2	450	2	A38316	alpha-2-adrenergic	338	26	81.2	710	1	C43497	early transcriptio
266	26	81.2	450	2	I49481	alpha-2-adrenergic	339	26	81.2	710	1	I42517	early transcriptio
267	26	81.2	450	2	B40392	alpha-2-adrenergic	340	26	81.2	710	2	T37394	VEGF, 82K subunit
268	26	81.2	457	2	T35443	aminotransferase -	341	26	81.2	714	2	G71963	fumarate reductase
269	26	81.2	457	2	S03961	Ig mu chain C regi	342	26	81.2	714	2	H64543	arginine decarboxy
270	26	81.2	459	2	G86212	hypothetical prote	343	26	81.2	728	1	S59553	transcription regu
271	26	81.2	460	2	JC5137	beta-glucosidase (344	26	81.2	729	2	C69038	melanotransferrin
272	26	81.2	460	2	F88085	protein Tf1p1.6 [i	345	26	81.2	738	1	TFHUM	hypothetical prote
273	26	81.2	462	2	AH1130	glutamate decarbox	346	26	81.2	738	2	T27289	hypothetical prote
274	26	81.2	463	2	T32111	hypothetical prote	347	26	81.2	740	2	T12223	NADH2 dehydrogenas
275	26	81.2	468	2	T10595	hypothetical prote	348	26	81.2	753	2	S48261	hypothetical prote
276	26	81.2	469	2	E70076	arabinan endo-1,5-	349	26	81.2	763	2	I50807	complement factor
277	26	81.2	475	2	T18487	hypothetical prote	350	26	81.2	764	1	BBHU	complement factor
278	26	81.2	475	2	A48948	aspartate carbamoy	351	26	81.2	778	2	B70667	hypothetical prote
279	26	81.2	476	2	E97631	probable membrane	352	26	81.2	783	2	A41627	furin (EC 3.4.21.7
280	26	81.2	485	2	T04840	hypothetical prote	353	26	81.2	794	2	T23048	hypothetical prote
281	26	81.2	488	2	D82991	conserved hypothet	354	26	81.2	817	2	T43659	trehalase-phosphat
282	26	81.2	489	2	C87480	conserved hypothet	355	26	81.2	822	2	S77112	DNA mismatch repai
283	26	81.2	489	2	F84725	probable protein k	356	26	81.2	826	2	T43638	casepase-related pr
284	26	81.2	489	2	F65125	probable general s	357	26	81.2	830	2	T04848	protein kinase hom
285	26	81.2	490	2	S28355	hypothetical prote	358	26	81.2	841	2	T14398	S-receptor kinase
286	26	81.2	501	2	T37747	hypothetical prote	359	26	81.2	849	2	T38728	probable trehalase
287	26	81.2	510	2	S47116	cellulose 1,4-beta	360	26	81.2	861	2	T15903	protein kinase C h
288	26	81.2	511	2	B87484	cellulose 1,4-beta	361	26	81.2	878	2	H83184	phosphoenolpyruvat
289	26	81.2	513	2	S44716	anthranilate synth	362	26	81.2	888	2	H88085	protein Tf1p1.8 [i
290	26	81.2	514	1	S21125	steroid 17alpha-mo	363	26	81.2	938	2	A56731	chromatin assembly
291	26	81.2	514	2	D84723	hypothetical prote	364	26	81.2	957	2	F69729	excinuclease ABC c
292	26	81.2	522	2	S33029	hypothetical prote	365	26	81.2	957	2	B84099	excinuclease ABC (
293	26	81.2	540	2	T41942	cellulose 1,4-beta	366	26	81.2	1004	2	S51133	transposase Tn4652
294	26	81.2	548	2	T28910	hypothetical prote	367	26	81.2	1023	2	T30257	IGG FC binding pro
295	26	81.2	550	2	G70768	hypothetical prote	368	26	81.2	1076	2	T26044	hypothetical prote
296	26	81.2	558	2	S68981	aminonucleoside an	369	26	81.2	1088	2	D82246	probable chitinase
297	26	81.2	566	2	F84903	probable ubiquitin	370	26	81.2	1137	2	T02097	glutamate synthase
298	26	81.2	569	2	T50711	urease (EC 3.5.1.5	371	26	81.2	1166	1	S06142	protein-tyrosine k
299	26	81.2	574	2	B29677	complement C9 prec	372	26	81.2	1185	2	F58009	gene ERCC5 protein
300	26	81.2	576	2	T41587	probable carbon ca	373	26	81.2	1186	2	S35993	DNA repair protein
301	26	81.2	583	2	S46265	cysteine proteinas	374	26	81.2	1192	2	S69000	laminin gamma 2 ch
302	26	81.2	588	2	I37202	B-CAM protein - hu	375	26	81.2	1240	2	T04833	hypothetical prote
303	26	81.2	596	2	B41627	furin (EC 3.4.21.7	376	26	81.2	1260	2	T01334	hypothetical prote
304	26	81.2	597	2	E85090	probable transposo	377	26	81.2	1270	2	T21269	hypothetical prote
305	26	81.2	597	2	T04954	hypothetical prote	378	26	81.2	1275	2	T18556	O-antigen biosynth
306	26	81.2	598	1	QBEG57	HHLF7 protein - hu	379	26	81.2	1318	2	T21266	hypothetical prote
307	26	81.2	609	1	FP00	alpha-fetoprotein	380	26	81.2	1327	2	T21268	hypothetical prote
308	26	81.2	609	1	FP00	alpha-fetoprotein	381	26	81.2	1331	2	S05011	calcium channel al
309	26	81.2	609	2	JC4258	alpha-fetoprotein	382	26	81.2	1381	2	S55619	capsid protein 25
310	26	81.2	612	2	T02414	probable protein k	383	26	81.2	1465	2	S31262	tyb protein - yeas
311	26	81.2	619	2	G84489	probable TNP2-like	384	26	81.2	1467	2	FC1253	tyb protein - yeas
312	26	81.2	628	2	I38000	Lutheran blood gro	385	26	81.2	1482	2	S67496	glutamate synthase
313	26	81.2	631	2	C96768	hypothetical prote	386	26	81.2	1531	2	T42218	glut-1 protein hom
314	26	81.2	633	2	T04835	probable serine/th	387	26	81.2	1538	2	S73296	glutamate synthase
315	26	81.2	635	2	B72215	hypothetical prote	388	26	81.2	1549	2	T11974	glutamate synthase
316	26	81.2	640	2	T06419	gag-proteinase pol	389	26	81.2	1556	2	S76781	glutamate synthase
317	26	81.2	650	2	A69390	3-hydroxyacyl-CoA	390	26	81.2	1559	2	A12348	ferredoxin-glutama
318	26	81.2	656	1	B44954	fumarate reductase	391	26	81.2	1580	2	T26204	hypothetical prote
319	26	81.2	658	2	T96656	hypothetical prote	392	26	81.2	1616	2	A38596	glutamate synthase
320	26	81.2	658	2	T04831	probable serine/th	393	26	81.2	1629	2	C84839	ferredoxin-depende
321	26	81.2	659	2	A64139	ribonuclease T2 ho	394	26	81.2	1802	2	S52611	tyb protein - yeas

395	26	81.2	1803	2	S56894	TYB protein - Yeas	468	25	78.1	276	2	F89826	hypothetical prote
396	26	81.2	1918	2	S43719	lactase (EC 3.2.1.1	469	25	78.1	278	2	T35146	glutamate-binding
397	26	81.2	1920	2	S43721	lactase (EC 3.2.1.1	470	25	78.1	281	2	T52470	cytochrome-b5 redu
398	26	81.2	1926	2	S01159	beta-glycosidase c	471	25	78.1	293	2	D96605	unknown protein F1
399	26	81.2	2051	2	T30938	receptor tyrosine	472	25	78.1	300	2	D95971	conserved hypochet
400	26	81.2	2108	2	S72458	sodium channel pro	473	25	78.1	300	2	AF3092	conserved hypochet
401	26	81.2	2166	2	S11339	calcium channel pr	474	25	78.1	300	2	D98194	hypothetical prote
402	26	81.2	2171	2	S05054	calcium channel al	475	25	78.1	303	2	A84831	hypothetical prote
403	26	81.2	2204	1	RRNZNV	genome polypeptin	476	25	78.1	304	2	E95342	NosF copper ABC tr
404	26	81.2	2267	2	T18497	hypothetical prote	477	25	78.1	308	2	B83403	ribokinase PA1950
405	26	81.2	2796	2	JC4743	fatty-acid synthas	478	25	78.1	318	1	LUP04	annexin IV - pig
406	26	81.2	2911	2	T20566	hypothetical prote	479	25	78.1	319	1	LUBO4	annexin IV - bovin
407	26	81.2	2924	2	T18378	variant-specific s	480	25	78.1	319	2	AC0942	annexin IV - human
408	26	81.2	3026	2	T28431	variant surface pr	481	25	78.1	321	1	A42077	annexin V - chicke
409	26	81.2	3028	2	T30835	breast cancer tumo	482	25	78.1	321	1	LUCH5	annexin V - chicke
410	26	81.2	3329	2	T42205	breast cancer susc	483	25	78.1	323	1	LUDU3	annexin III - huma
411	26	81.2	3329	2	T30904	breast cancer tumo	484	25	78.1	323	2	A70404	acetyl-CoA carboxy
412	26	81.2	3430	1	GNWVW	genome polypeptin	485	25	78.1	323	1	LUR73	annexin III - rat
413	26	81.2	3432	1	GNWVJ	genome polypeptin	486	25	78.1	325	2	T04677	hypothetical prote
414	26	81.2	3432	1	GNWVJ	genome polypeptin	487	25	78.1	325	2	AE2466	acetyl-CoA carboxy
415	26	81.2	3433	1	GNWVK	genome polypeptin	488	25	78.1	326	2	S40818	hypothetical 36.9K
416	26	81.2	3434	1	GNWVW	genome polypeptin	489	25	78.1	326	2	S40818	hypothetical 36.9K
417	26	81.2	4131	2	T21085	hypothetical prote	490	25	78.1	335	2	G01448	homeotic protein E
418	26	81.2	5147	1	IJFF7M	cadherin-related t	491	25	78.1	339	1	LUBO36	homeobox protein H
419	25	78.1	36	2	A60602	annexin II - golde	492	25	78.1	339	1	LUDU36	annexin II - bovin
420	25	78.1	85	2	T10571	hypothetical prote	493	25	78.1	339	1	LUMS36	annexin II - mouse
421	25	78.1	89	2	B91173	phosphotransferase	494	25	78.1	339	2	S33700	annexin II - rat
422	25	78.1	89	2	B86019	probable phosphoca	495	25	78.1	340	1	JQ1297	annexin II type 1
423	25	78.1	109	2	B72763	hypothetical prote	496	25	78.1	340	1	JQ1298	annexin II type 2
424	25	78.1	110	2	E85547	hypothetical prote	497	25	78.1	341	2	S55277	annexin II - rat
425	25	78.1	113	2	H86934	probable transcrip	498	25	78.1	352	2	AI2884	probable ABC trans
426	25	78.1	120	1	PSKF3U	probable membrane	499	25	78.1	352	2	F97660	probable ABC trans
427	25	78.1	122	1	PSABA	phospholipase A2 (500	25	78.1	353	2	S33322	(S)-2-hydroxy-acid
428	25	78.1	122	1	PSRSE	phospholipase A2 (501	25	78.1	353	2	T42517	bromodomain protei
429	25	78.1	122	1	PSRSAW	phospholipase A2 (502	25	78.1	363	2	T06726	cysteine proteinas
430	25	78.1	122	2	A33317	phospholipase A2 (503	25	78.1	368	2	T46607	3-isopropylmalate
431	25	78.1	122	2	A39557	phospholipase A2 (504	25	78.1	387	2	JL0101	methane monooxygen
432	25	78.1	122	2	B37478	phospholipase A2 (505	25	78.1	396	2	E96911	glutamyl-tRNA redu
433	25	78.1	122	2	A37478	phospholipase A2 (506	25	78.1	397	2	E69173	protoporphyrin IX
434	25	78.1	122	2	A44179	phospholipase A2 (507	25	78.1	401	2	E96570	hypothetical prote
435	25	78.1	124	2	A26535	phospholipase A2 (508	25	78.1	403	2	T25556	hypothetical prote
436	25	78.1	124	2	A59420	phospholipase A2 (509	25	78.1	407	2	C36786	hypothetical prote
437	25	78.1	125	2	A45660	lysozyme (EC 3.2.1	510	25	78.1	412	2	C38104	LFY floral meriste
438	25	78.1	129	2	S07290	lysozyme (EC 3.2.1	511	25	78.1	420	2	B38104	LFY floral meriste
439	25	78.1	129	2	G35558	lysozyme (EC 3.2.1	512	25	78.1	423	2	E84751	hypothetical prote
440	25	78.1	132	2	I39193	gene HOXA1 protein	513	25	78.1	424	2	AI3606	adenosylmethionine
441	25	78.1	133	2	B30242	stem cell protein	514	25	78.1	424	2	A38104	LFY floral meriste
442	25	78.1	135	2	F64779	probable transcrip	515	25	78.1	424	2	T35535	probable secreted
443	25	78.1	135	2	A90697	probable transcrip	516	25	78.1	429	2	S72829	trehalose-6-phosph
444	25	78.1	137	1	JC4877	phospholipase A2 h	517	25	78.1	435	2	D81272	hypothetical prote
445	25	78.1	137	2	S22388	phospholipase A2 (518	25	78.1	456	2	S55661	hypothetical prote
446	25	78.1	138	1	I51380	phospholipase A2 h	519	25	78.1	465	2	T00090	glycosyltransferas
447	25	78.1	138	2	JC1342	phospholipase A2 (520	25	78.1	470	2	S54089	hypothetical prote
448	25	78.1	138	2	AF0564	copper efflux regu	521	25	78.1	471	2	T04738	hypothetical prote
449	25	78.1	147	1	LZDK	lysozyme (EC 3.2.1	522	25	78.1	476	2	T35528	probable aminotran
450	25	78.1	155	2	C64314	conserved hypochet	523	25	78.1	478	2	F89651	protein T04P8.2 [i
451	25	78.1	155	2	C69002	conserved hypochet	524	25	78.1	484	2	T50909	hypothetical prote
452	25	78.1	167	2	T47385	hypothetical prote	525	25	78.1	491	2	T21421	hypothetical prote
453	25	78.1	185	2	H87107	ribosome recycling	526	25	78.1	494	2	C70940	probable cobQ prot
454	25	78.1	186	2	S65723	extracellular hemo	527	25	78.1	499	2	T49064	cytochrome P450-li
455	25	78.1	201	2	S52833	probable heme bind	528	25	78.1	502	2	T24471	hypothetical prote
456	25	78.1	204	2	B84173	hypothetical prote	529	25	78.1	503	2	T70853	probable Zn finger
457	25	78.1	211	2	G84202	hypothetical prote	530	25	78.1	509	2	T02864	probable Zn finger
458	25	78.1	217	2	T48662	hypothetical prote	531	25	78.1	541	2	S56653	thioglucoosidase (E
459	25	78.1	218	2	G98291	hypothetical prote	532	25	78.1	550	1	KJHUID	iduronate-2-sulfat
460	25	78.1	219	2	JQ1692	aspartic proteinas	533	25	78.1	607	2	F86230	hypothetical prote
461	25	78.1	221	2	T07411	proteinase inhibit	534	25	78.1	621	2	I38467	low density lipopr
462	25	78.1	221	2	T07413	proteinase inhibit	535	25	78.1	632	2	T32454	hypothetical prote
463	25	78.1	222	2	I39192	gene HOXA1 protein	536	25	78.1	636	2	T37843	probable beta-gluc
464	25	78.1	231	2	H84535	hypothetical prote	537	25	78.1	640	2	T51231	R-protein promoter
465	25	78.1	264	1	A53239	aminoglycoside-O-p	538	25	78.1	646	2	JQ2150	B west mating prot
466	25	78.1	271	2	AF1157	phosphomethylpyrim	539	25	78.1	666	2	D69366	conserved hypochet
467	25	78.1	271	2	AC1516	phosphomethylpyrim	540	25	78.1	673	1	S01786	annexin VI - mouse

541	25	78.1	673	1	S52844	annexin VI - rat	614	24	75.0	182	2	F97201	hypothetical prote
542	25	78.1	694	2	D87330	sensory box/GGDEF	615	24	75.0	183	2	S05569	signaling protein
543	25	78.1	712	2	T02470	potassium transpor	616	24	75.0	183	2	B90643	Ampr protein [limp
544	25	78.1	743	2	E71432	hypothetical prote	617	24	75.0	183	2	B85494	regulates ampc lim
545	25	78.1	754	2	T45910	hypothetical prote	618	24	75.0	183	2	F69049	conserved ampc lim
546	25	78.1	768	2	T17462	disease resistance	619	24	75.0	184	2	A82371	carbonic anhydrase
547	25	78.1	782	2	I48746	senaphorin C - mou	620	24	75.0	187	2	A10520	Ampr protein (anhy
548	25	78.1	784	2	T26585	hypothetical prote	621	24	75.0	191	2	AC3285	leucine-, isoleuci
549	25	78.1	792	2	G84830	probable potassium	622	24	75.0	206	2	D70628	hypothetical prote
550	25	78.1	797	2	T22294	hypothetical prote	623	24	75.0	210	2	AC0259	hypothetical phage
551	25	78.1	843	2	T01438	hypothetical prote	624	24	75.0	224	2	B96925	hypothetical prote
552	25	78.1	854	1	ORHYLD	LDL receptor precu	625	24	75.0	229	2	T09215	H+-exporting ATPas
553	25	78.1	855	2	T17460	disease resistance	626	24	75.0	230	2	I48685	mast cell proteina
554	25	78.1	855	2	T07015	Cf-4A protein - to	627	24	75.0	242	2	T40458	hypothetical prote
555	25	78.1	860	2	T16892	hypothetical prote	628	24	75.0	244	2	S75653	hypothetical prote
556	25	78.1	881	2	F84404	valyl-tRNA synthet	629	24	75.0	245	2	G82175	hypothetical prote
557	25	78.1	896	2	S48761	trehalose-phosphat	630	24	75.0	252	2	B69034	hypothetical prote
558	25	78.1	931	2	AF3276	ATP-dependent clip	631	24	75.0	256	2	T21238	hypothetical prote
559	25	78.1	956	1	A46016	thrombospondin 3 -	632	24	75.0	262	2	T51409	MADS box protein A
560	25	78.1	956	2	A57121	thrombospondin 3 p	633	24	75.0	272	2	T36770	probable expressio
561	25	78.1	969	2	H69425	modification methy	634	24	75.0	284	2	F82666	phage-related prot
562	25	78.1	996	2	JB0237	apolipoprotein E r	635	24	75.0	293	2	S75355	probable [acyl-car
563	25	78.1	1027	2	A56533	chromatin remodel	636	24	75.0	296	2	D36788	hypothetical prote
564	25	78.1	1046	2	F71432	hypothetical prote	637	24	75.0	298	2	A49630	ubiquitin conjugat
565	25	78.1	1049	2	JC4783	sucrose-phosphate	638	24	75.0	301	2	E71482	phosphatidylserine
566	25	78.1	1084	2	T04103	sucrose-phosphate	639	24	75.0	305	2	T27930	hypothetical prote
567	25	78.1	1160	2	F88369	protein unc-52 lim	640	24	75.0	306	2	D82497	hypothetical prote
568	25	78.1	1237	2	E86457	probable RNA helic	641	24	75.0	308	2	A10001	ribokinase IEC 2.7
569	25	78.1	1254	1	A32686	DNA-directed DNA p	642	24	75.0	308	2	T26716	hypothetical prote
570	25	78.1	1268	2	T31420	C-terminal domain-	643	24	75.0	310	2	T30152	hypothetical prote
571	25	78.1	1347	2	F84531	probable retroelem	644	24	75.0	310	2	T43147	hypothetical prote
572	25	78.1	1358	2	F84531	genome polypotein	645	24	75.0	312	2	S75932	hypothetical prote
573	25	78.1	1444	2	T18856	angiogenesis inhib	646	24	75.0	317	2	C97245	probable phosphohy
574	25	78.1	1451	1	JQ1719	E2 glycoprotein pr	647	24	75.0	318	2	T39594	ribokinase - fisci
575	25	78.1	1453	2	S41453	spike protein - ca	648	24	75.0	324	2	B82741	thiamin monophosph
576	25	78.1	1460	1	EDBE1P	immediate-early pr	649	24	75.0	326	2	T51941	homocysteine S-met
577	25	78.1	1468	2	S11515	formin - mouse	650	24	75.0	332	2	S44743	C02D5.1 protein -
578	25	78.1	1694	2	JQ1896	hypothetical 191.1	651	24	75.0	343	2	C85020	hypothetical prote
579	25	78.1	2165	2	T21371	hypothetical prote	652	24	75.0	343	2	T03960	r40g2 protein - ri
580	25	78.1	2227	1	GNNYHB	genome polypotein	653	24	75.0	349	2	AF2967	trehalose utilizat
581	25	78.1	2227	1	GNNYHB	genome polypotein	654	24	75.0	349	2	E98315	thub protein (AF17
582	25	78.1	2227	1	GNNYHR	genome polypotein	655	24	75.0	351	2	T22861	hypothetical prote
583	25	78.1	2227	1	GNNYHR	genome polypotein	656	24	75.0	354	1	VWVPLJ	coat protein vpi -
584	25	78.1	2228	2	T14029	variant-specific s	657	24	75.0	357	2	S56601	hypothetical 39.8K
585	25	78.1	2230	1	GNNYSA	genome polypotein	658	24	75.0	357	2	B86137	hypothetical prote
586	25	78.1	2295	2	C88369	protein unc-52 lim	659	24	75.0	357	2	G91295	hypothetical prote
587	25	78.1	2643	2	T29149	hypothetical prote	660	24	75.0	359	2	AH3465	oxidoreductase IEC
588	25	78.1	2764	2	T13949	neurofibromin - fr	661	24	75.0	361	2	T48014	serine/threonine p
589	25	78.1	2802	2	T13945	neurofibromin - fr	662	24	75.0	362	2	E95972	probable sugar upt
590	25	78.1	2802	2	T13947	neurofibromin - fr	663	24	75.0	363	2	C87336	serine proteinase
591	25	78.1	2895	2	H85362	hypothetical prote	664	24	75.0	374	1	A56436	alcohol dehydrogen
592	25	78.1	3078	2	T28432	variant-specific s	665	24	75.0	374	1	DEHUAS	alcohol dehydrogen
593	25	78.1	3375	2	T19821	hypothetical prote	666	24	75.0	375	1	A33909	alcohol dehydrogen
594	25	78.1	3573	2	S23070	erythronolide synt	667	24	75.0	375	1	DEHUAB	hypothetical prote
595	25	78.1	4660	2	T42737	gpi30 protein prec	668	24	75.0	380	2	T19096	hypothetical prote
596	25	78.1	4753	1	A47437	LDL-receptor-relat	669	24	75.0	382	2	S50762	hydroxymethylbilan
597	24	75.0	54	2	H81814	probable small sec	670	24	75.0	412	1	G69965	ATP/GTP-binding pr
598	24	75.0	60	2	A54369	fibrinogen recepto	671	24	75.0	425	2	A31026	probable membrane
599	24	75.0	62	2	S01803	chymotrypsin inhib	672	24	75.0	427	1	XUECVS	3-phosphoshikimate
600	24	75.0	73	2	T31219	hypothetical prote	673	24	75.0	427	2	E85616	5-enolpyruvylshiki
601	24	75.0	112	2	A47130	neurotoxin txi pre	674	24	75.0	427	2	G90752	5-enolpyruvylshiki
602	24	75.0	119	2	JN0403	phospholipase A2 (675	24	75.0	428	2	S41485	sp2 protein - mous
603	24	75.0	119	2	S07528	phospholipase A2 (676	24	75.0	432	2	AD0085	probable type II s
604	24	75.0	149	2	C95917	hypothetical prote	677	24	75.0	432	2	T43476	hypothetical prote
605	24	75.0	150	2	G83840	ferric uptake regu	678	24	75.0	434	2	T01938	hypothetical prote
606	24	75.0	160	2	F82664	hypothetical prote	679	24	75.0	438	2	H70844	probable aminoran
607	24	75.0	164	2	B4716	conserved hypothet	680	24	75.0	442	2	AG2325	hypothetical prote
608	24	75.0	164	2	G71802	hypothetical prote	681	24	75.0	443	2	S56609	Yjfy protein - Esc
609	24	75.0	170	2	AG2479	aminoglycoside N6,	682	24	75.0	446	2	S69051	hypothetical prote
610	24	75.0	171	1	B64671	serine O-acetyltra	683	24	75.0	462	2	G69220	sensory transducti
611	24	75.0	171	2	B71845	o-serine acetyltra	684	24	75.0	462	2	A57120	small nuclear ribo
612	24	75.0	172	2	F81439	translation initia	685	24	75.0	478	2	F96842	hypothetical prote
613	24	75.0	173	2	D84564	hypothetical prote	686	24	75.0	481	2	B64585	cag pathogenicity

687	24	75.0	483	2	I40031	uoporphyrin-III C	760	24	75.0	1429	2	S16233	nitric-oxide synth
688	24	75.0	486	2	S43274	gag-like protein -	761	24	75.0	1433	2	G01946	nitric-oxide synth
689	24	75.0	497	2	B97555	hypothetical prote	762	24	75.0	1436	2	A46496	antigen WCl_1 prec
690	24	75.0	497	2	AD2775	fumarate hydratase	763	24	75.0	1484	2	C97196	probable membrane
691	24	75.0	502	2	T02306	probable protein k	764	24	75.0	1584	2	T22674	hypothetical prote
692	24	75.0	510	2	F96649	hypothetical prote	765	24	75.0	2214	2	T16305	hypothetical prote
693	24	75.0	533	2	T49061	hypothetical prote	766	24	75.0	2254	2	T09053	low voltage-activa
694	24	75.0	541	2	A54841	DFT1 protein - Ara	767	24	75.0	4574	2	G02520	plectin - human
695	24	75.0	552	2	C86171	hypothetical prote	768	24	75.0	4684	2	A59404	plectin - human
696	24	75.0	557	2	T16696	hypothetical prote	769	24	75.0	4687	1	A39638	plectin [imported]
697	24	75.0	557	2	A48434	variant-specific s	770	23	71.9	36	2	C82694	hypothetical prote
698	24	75.0	581	2	B87678	hypothetical prote	771	23	71.9	41	2	S19566	ornatin A2 - leech
699	24	75.0	594	2	A86309	F20D23.3 protein -	772	23	71.9	41	2	S19621	ornatin A3 - leech
700	24	75.0	597	2	T16006	protein kinase C3	773	23	71.9	47	2	A53982	echistatin alpha-2
701	24	75.0	599	2	G71481	probable phosphoen	774	23	71.9	49	2	S53432	echistatin gamma -
702	24	75.0	600	2	E72027	phosphoenolpyruvat	775	23	71.9	49	2	A32029	echistatin alpha-1
703	24	75.0	600	2	A86597	phosphoenolpyruvat	776	23	71.9	50	2	A69055	hypothetical prote
704	24	75.0	607	2	T45969	mRNA capping enzym	777	23	71.9	55	1	BGHS	spermatid transiti
705	24	75.0	608	2	JQ1462	phosphoenolpyruvat	778	23	71.9	55	1	BGHS	spermatid transiti
706	24	75.0	619	2	A45625	phosphoenolpyruvat	779	23	71.9	55	1	BGHS	spermatid transiti
707	24	75.0	622	1	QYCHGC	phosphoenolpyruvat	780	23	71.9	60	1	T6BP3D	short toxin 3 - bl
708	24	75.0	622	1	QYRTCP	phosphoenolpyruvat	781	23	71.9	60	1	A39165	calcisepine - bla
709	24	75.0	622	2	A45746	phosphoenolpyruvat	782	23	71.9	62	1	DNVPA4	DNA-binding protei
710	24	75.0	631	2	T38167	electron transfer	783	23	71.9	62	2	S01802	chymotrypsin inhib
711	24	75.0	634	2	T17232	hypothetical prote	784	23	71.9	63	1	TIIMTC3	chymotrypsin inhib
712	24	75.0	640	1	QYCHQM	phosphoenolpyruvat	785	23	71.9	66	2	T36227	hypothetical prote
713	24	75.0	640	2	S69546	phosphoenolpyruvat	786	23	71.9	71	2	A59409	platelet-aggregati
714	24	75.0	640	2	T08758	hypothetical prote	787	23	71.9	71	2	A59411	platelet-aggregati
715	24	75.0	647	1	QIFFQM	phosphoenolpyruvat	788	23	71.9	71	2	A33990	apilaggin - easter
716	24	75.0	653	2	H97777	DNA helicase II (E	789	23	71.9	73	2	A59410	platelet aggregati
717	24	75.0	654	2	T24168	hypothetical prote	790	23	71.9	73	2	A23731	albolabrin - green
718	24	75.0	656	2	JC2005	integrin beta-5 ch	791	23	71.9	73	2	E35982	trigramin gamma -
719	24	75.0	656	2	B84547	hypothetical prote	792	23	71.9	73	2	D35982	trigramin gamma -
720	24	75.0	658	2	F71703	DNA helicase II (u	793	23	71.9	75	2	B45206	trigramin beta-2 -
721	24	75.0	659	2	F70175	rep helicase, sing	794	23	71.9	75	2	T16742	metallothionein 2
722	24	75.0	660	2	T20569	hypothetical prote	795	23	71.9	77	2	G72214	hypothetical prote
723	24	75.0	680	2	G82526	ATP-dependent DNA	796	23	71.9	81	2	AG0590	hypothetical prote
724	24	75.0	713	2	B38992	cadherin 13 precu	797	23	71.9	84	2	B82599	hypothetical prote
725	24	75.0	727	2	A56879	diacylglycerol kin	798	23	71.9	85	1	IMECE8	colicin E8 immunit
726	24	75.0	735	2	AC1946	adenylate cyclase	799	23	71.9	85	2	I40591	immunity protein 8
727	24	75.0	759	2	D69846	ATP-dependent DNA	800	23	71.9	85	2	JQ0327	colicin E8 immunit
728	24	75.0	775	2	C81594	hypothetical prote	801	23	71.9	85	2	C87181	hypothetical prote
729	24	75.0	775	2	B86549	hypothetical prote	802	23	71.9	86	2	A46461	T-cell receptor et
730	24	75.0	775	2	B72074	hypothetical prote	803	23	71.9	86	2	S38920	hypothetical prote
731	24	75.0	792	2	S32244	x-pro dipeptidyl-p	804	23	71.9	88	2	T06986	hypothetical prote
732	24	75.0	795	2	B97294	stage II sporulati	805	23	71.9	89	2	T43992	hypothetical prote
733	24	75.0	796	2	T34805	hypothetical prote	806	23	71.9	93	2	T35660	probable acylphosp
734	24	75.0	799	2	A38308	integrin beta-5 ch	807	23	71.9	97	2	C81102	conserved hypotet
735	24	75.0	805	2	G87438	DNA helicase II [i	808	23	71.9	99	2	H72868	AcOrf-150 protein
736	24	75.0	824	2	AB2829	DNA helicase II [i	809	23	71.9	102	2	S72839	probable transcrip
737	24	75.0	826	2	T01494	trehalose-6-phosph	810	23	71.9	102	2	E70737	probable whiB3 pro
738	24	75.0	827	2	H97606	probable DNA helic	811	23	71.9	104	2	S15395	tissue kallikrein-
739	24	75.0	827	2	AF3326	DNA helicase II (E	812	23	71.9	105	2	T49850	hypothetical prote
740	24	75.0	847	2	C96703	hypothetical prote	813	23	71.9	107	2	S60590	Ig heavy chain var
741	24	75.0	851	2	A86200	hypothetical prote	814	23	71.9	108	1	PVLAB	parvalbumin beta -
742	24	75.0	853	2	S49876	gamma-adaptin - sm	815	23	71.9	109	2	S58825	hypothetical prote
743	24	75.0	861	2	T02267	trehalose-6-phosph	816	23	71.9	112	2	B69468	nitrogen regulator
744	24	75.0	862	2	E84567	probable trehalose	817	23	71.9	112	2	H70232	hypothetical prote
745	24	75.0	865	2	H71447	trehalose-6-phosph	818	23	71.9	112	2	T35596	probable regulator
746	24	75.0	870	2	T31795	hypothetical prote	819	23	71.9	112	2	H75064	tetracenomyacin pol
747	24	75.0	884	2	S53396	telomerase catalyt	820	23	71.9	114	2	I50606	brain-derived neur
748	24	75.0	919	2	T32541	unc-5 protein - Ca	821	23	71.9	114	2	I84765	brain-derived neur
749	24	75.0	943	2	G84402	dipeptide ABC tran	822	23	71.9	114	2	B75060	hypothetical prote
750	24	75.0	943	2	S59317	DIP2 protein - yea	823	23	71.9	115	2	C86412	protein Flk23.7 [i
751	24	75.0	947	1	B44294	unc-5 protein, lon	824	23	71.9	116	1	NGNXXI	nerve growth facto
752	24	75.0	966	1	FLVBEB	RNA Ia protein - b	825	23	71.9	116	2	A58566	nerve growth facto
753	24	75.0	980	2	T05414	protein kinase hom	826	23	71.9	116	2	D84285	hypothetical prote
754	24	75.0	983	2	T19874	hypothetical prote	827	23	71.9	117	2	S28161	nerve growth facto
755	24	75.0	1032	2	T20342	hypothetical prote	828	23	71.9	118	2	G84322	50S ribosomal prot
756	24	75.0	1127	2	D70671	pyruvate carboxyla	829	23	71.9	119	2	A39426	saxiphilin - bullf
757	24	75.0	1142	2	T30272	hypothetical prote	830	23	71.9	121	1	A37331	secretory protein
758	24	75.0	1250	2	T00454	hypothetical prote	831	23	71.9	121	2	H72507	hypothetical prote
759	24	75.0	1429	2	JN0609	nitric-oxide synth	832	23	71.9	124	2	A81657	conserved hypotet

833	23	71.9	125	2	A26312	nerve growth facto	306	23	71.9	180	2	C81664	ribosomal protein
834	23	71.9	127	2	B84172	hypothetical prote	307	23	71.9	180	2	E71506	ribosomal protein
835	23	71.9	128	1	I64059	mercuric resistanc	308	23	71.9	180	2	T07702	hypothetical prote
836	23	71.9	128	2	T06028	hypothetical prote	309	23	71.9	180	2	AC3270	hypothetical expor
837	23	71.9	128	2	AB2682	carboxymuconolacto	310	23	71.9	180	2	D87474	conserved hypothet
838	23	71.9	130	2	S66474	cytochrome P450 (C	311	23	71.9	181	2	A64393	hypothetical prote
839	23	71.9	131	2	S66475	cytochrome P450 (C	312	23	71.9	181	2	S77302	hypothetical prote
840	23	71.9	131	2	S23511	thionin type V pre	313	23	71.9	181	2	D69129	conserved hypothet
841	23	71.9	131	2	B75350	hypothetical prote	314	23	71.9	183	2	S36038	BHUH m8 protein -
842	23	71.9	131	2	F70880	hypothetical prote	315	23	71.9	183	2	C83913	hypothetical prote
843	23	71.9	132	2	T48370	hypothetical prote	316	23	71.9	184	2	T13186	hypothetical prote
844	23	71.9	132	2	E82518	hypothetical prote	317	23	71.9	184	2	B87708	hypothetical prote
845	23	71.9	132	2	D83003	hypothetical prote	318	23	71.9	185	1	HLMS2	MHC class I histoc
846	23	71.9	134	2	S66694	histone H2A.HTA3 -	319	23	71.9	186	2	E69537	conserved hypothet
847	23	71.9	135	2	C70890	hypothetical prote	320	23	71.9	187	2	D72501	hypothetical prote
848	23	71.9	136	2	F64416	hypothetical prote	321	23	71.9	187	2	H75606	hypothetical prote
849	23	71.9	137	2	F69436	conserved hypothet	322	23	71.9	190	2	JC4514	TATA-binding prote
850	23	71.9	138	2	B90665	hypothetical prote	323	23	71.9	190	2	H71245	hypothetical prote
851	23	71.9	138	2	B85515	unknown protein en	324	23	71.9	191	2	S54115	complement factor
852	23	71.9	138	2	AH1214	hypothetical prote	325	23	71.9	191	2	E75072	transcription init
853	23	71.9	139	2	E71110	hypothetical prote	326	23	71.9	191	2	S56012	XS-2 protein (homo
854	23	71.9	140	2	F69010	hypothetical prote	327	23	71.9	191	2	A81870	hypothetical prote
855	23	71.9	141	2	G87702	conserved hypothet	328	23	71.9	191	2	D72711	hypothetical prote
856	23	71.9	146	2	A25233	replication protei	329	23	71.9	191	2	D71093	probable TATA-bind
857	23	71.9	146	2	H84259	hypothetical prote	330	23	71.9	193	2	T35847	probable carbonic
858	23	71.9	147	2	C84318	hypothetical prote	331	23	71.9	193	2	PQ0503	surface protein -
859	23	71.9	148	2	B70964	hypothetical prote	332	23	71.9	193	2	PQ0504	surface protein -
860	23	71.9	149	2	AG1991	hypothetical prote	333	23	71.9	193	2	H64174	hypothetical prote
861	23	71.9	150	2	C82994	conserved hypothet	334	23	71.9	193	2	B64435	hypothetical prote
862	23	71.9	150	2	S15914	hypothetical prote	335	23	71.9	193	2	AH2528	hypothetical prote
863	23	71.9	153	2	H81413	hypothetical prote	336	23	71.9	195	2	E27733	nifQ protein - Azo
864	23	71.9	154	2	A60998	replication protei	337	23	71.9	195	2	T32819	hypothetical prote
865	23	71.9	154	2	B81810	hypothetical prote	338	23	71.9	195	2	A85364	hypothetical prote
866	23	71.9	156	2	T49921	ribosomal protein-1	339	23	71.9	197	2	S25825	polygalacturonase
867	23	71.9	156	2	S70970	bfgu protein - Esc	340	23	71.9	197	2	C83003	imidazoleglycerol-
868	23	71.9	156	2	A70968	hypothetical prote	341	23	71.9	200	2	F95985	conserved hypothet
869	23	71.9	157	1	EDXP3	immediate-early pr	342	23	71.9	201	2	S12789	GTP-binding protei
870	23	71.9	157	2	S77122	hypothetical prote	343	23	71.9	203	2	G84284	hypothetical prote
871	23	71.9	157	2	H89967	hypothetical prote	344	23	71.9	203	2	A10367	probable lipoprote
872	23	71.9	158	1	T05710	2S albumin precurs	345	23	71.9	205	2	E84334	hypothetical prote
873	23	71.9	158	1	YQBZN	fimbrial protein p	346	23	71.9	205	2	T08272	probable thioredox
874	23	71.9	158	2	A29827	replication protei	347	23	71.9	206	2	T08285	hypothetical prote
875	23	71.9	158	2	A51069	replication protei	348	23	71.9	206	2	T20376	hypothetical prote
876	23	71.9	158	2	T30791	hypothetical prote	349	23	71.9	206	2	E83047	hypothetical prote
877	23	71.9	158	2	B42508	F15L protein - vac	350	23	71.9	206	2	T33097	hypothetical prote
878	23	71.9	158	2	D82782	hypothetical prote	351	23	71.9	206	2	C82588	conserved hypothet
879	23	71.9	159	2	I84615	coagulation factor	352	23	71.9	207	2	A48608	E1 glycoprotein -
880	23	71.9	160	2	PL0105	anti-PR2 erythrocy	353	23	71.9	208	1	S26196	imidazoleglycerol-
881	23	71.9	160	2	F81844	hypothetical prote	354	23	71.9	208	2	A82077	MutT/nudix family
882	23	71.9	161	2	T28477	hypothetical prote	355	23	71.9	209	2	A83524	hypothetical prote
883	23	71.9	161	2	E72155	E15L protein - var	356	23	71.9	209	2	B42687	neurotrophin-4 pre
884	23	71.9	161	2	A36841	C19L protein - var	357	23	71.9	210	2	A42687	neurotrophin-4 pre
885	23	71.9	162	2	B87655	conserved hypothet	358	23	71.9	210	2	D85098	hypothetical prote
886	23	71.9	162	2	B24497	replication protei	359	23	71.9	210	2	T15105	hypothetical prote
887	23	71.9	162	2	E89011	3-isopropylmalate	360	23	71.9	212	2	A25856	ubiquitin thiolest
888	23	71.9	164	2	B84283	hypothetical prote	361	23	71.9	212	2	E86561	glycerol-3-P acylt
889	23	71.9	164	2	G84290	hypothetical prote	362	23	71.9	212	2	B72063	probable 1-acylgly
890	23	71.9	164	2	AC3046	hypothetical prote	363	23	71.9	212	2	H81344	hypothetical prote
891	23	71.9	164	2	S57074	HIT1 protein - yea	364	23	71.9	213	2	T03931	DNA binding protei
892	23	71.9	167	2	T43621	hypothetical prote	365	23	71.9	213	2	S50799	probable membrane
893	23	71.9	168	2	B64333	formate hydrogenly	366	23	71.9	214	2	H82314	conserved hypothet
894	23	71.9	168	2	B69059	acetolactate synth	367	23	71.9	214	2	A70593	probable tmk prote
895	23	71.9	169	2	AC0556	conserved hypothet	368	23	71.9	214	2	T19247	hypothetical prote
896	23	71.9	169	2	T15611	hypothetical proté	369	23	71.9	215	1	C69015	probable uracil ph
897	23	71.9	172	2	T39026	conserved hypothet	370	23	71.9	216	2	S49483	somatotropin precu
898	23	71.9	172	2	F84383	hypothetical prote	371	23	71.9	217	2	T31449	dheme cytochrome-
899	23	71.9	174	2	H97463	gamma-carboxymuon	372	23	71.9	219	2	F83160	two-component resp
900	23	71.9	175	2	T51077	hypothetical prote	373	23	71.9	219	2	F70779	hypothetical prote
901	23	71.9	175	2	D86699	hypothetical prote	374	23	71.9	221	2	T04238	hypothetical prote
902	23	71.9	177	2	T32766	hypothetical prote	375	23	71.9	222	2	E84283	TRK potassium upta
903	23	71.9	178	2	B81151	conserved hypothet	376	23	71.9	223	2	JX0222	ubiquitin thiolest
904	23	71.9	179	2	S03627	neurogenic gene co	377	23	71.9	223	2	A81068	hypothetical prote
905	23	71.9	179	2	T22521	hypothetical prote	378	23	71.9	225	2	AH0839	probable transcrip

979 23 71.9 225 2 T20325 hypothetical prote
980 23 71.9 226 2 D70540 probable dehiobio
981 23 71.9 226 2 T14675 hypothetical prote
982 23 71.9 227 2 T10510 cbbj protein - Rho
983 23 71.9 227 2 B71358 hypothetical prote
984 23 71.9 228 2 S28019 gene iofap protein
985 23 71.9 229 2 T46614 nerve growth facto
986 23 71.9 230 2 T51265 probable HLH DNA b
987 23 71.9 230 2 AH1864 hypothetical prote
988 23 71.9 231 2 A60468 venombin A (EC 3.4
989 23 71.9 231 2 A10372 hypothetical prote
990 23 71.9 232 2 A87363 dnaj family protei
991 23 71.9 232 2 T14939 hypothetical prote
992 23 71.9 233 2 D69181 hypothetical prote
993 23 71.9 234 2 T30427 probable apoptosis
994 23 71.9 235 1 B41326 nitrite hydratase
995 23 71.9 235 2 S14481 nerve growth facto
996 23 71.9 235 2 JC4603 conserved hypothet
997 23 71.9 236 2 JH0400 neurotrophin-4 pre
998 23 71.9 237 2 C95370 probable oxidoredu
999 23 71.9 238 2 AF3608 two component resp
1000 23 71.9 238 2 T11718 hypothetical prote

ALIGNMENTS

RESULT 1
G64320
C:Species: Methanococcus jannaschii
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C:Accession: G64320
R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Frazer, C.M.; Smith, H.O.; Woese, C
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A:Reference number: A64300; MUID:96337999; PMID:8688087
A:Accession: G64320
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-255 <BUL>
A:Cross-references: UNIPROT:Q57630; GB:U67473; GB:L77117; NID:g2826256; PIDN:AAB98148.1;
C:Genetics:
A:Map position: REV170194-169427
C:Superfamily: Pyrococcus horikoshii hypothetical protein PH0006

Query Match 100.0%; Score 32; DB 2; Length 255;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
|
|
|
|
Db 233 CRGDD 237

RESULT 2
T44599
C:Species: Rhizobium meliloti
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C:Accession: T44599
R;Kroll, E.A.; Soberon, M.; Yurgel, S.N.; Miranda Rios, J.; Simarov, B.V.
submitted to the EMBL Data Library, May 1999
A:Reference number: Z22806
A:Accession: T44599
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-289 <KRO>
A:Cross-references: UNIPROT:Q9X6W2; EMBL:AF148072; PIDN:AAD3993.1
A:Experimental source: strain CXM1-188

C:Genetics:
A:Gene: oppC
A:Map position: megaplasmid 2
A:Genome: plasmid
C:Superfamily: oligopeptide permease protein oppB
Query Match 100.0%; Score 32; DB 2; Length 289;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CRGDD 5
|
|
|
|
Db 36 CRGDD 40
RESULT 3
AG1072
C:Species: Salmonella enterica subsp. enterica serovar Typhi
C:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AG1072
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AG1072
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-357 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD03398.1; PID:g16505667; GSPDB:GN00176
C:Genetics:
A:Gene: STY4914
Query Match 100.0%; Score 32; DB 2; Length 357;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CRGDD 5
|
|
|
|
Db 139 CRGDD 143
RESULT 4
G95937
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C:Accession: G95937
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernanc
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1,683-kb pSymB megaplasamid from the N2-fixing endo
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: G95937
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-363 <KUR>
A:Cross-references: UNIPROT:Q92VE0; GB:AL591985; PIDN:CAC49167.1; PID:g15140652; GSPDB:GB
A:Experimental source: strain 1021, megaplasamid pSymB
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, P.; Barloy-Hubler, P.
Pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.C
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: SMB21263

A:Genome: plasmid
C:Superfamily: oligopeptide permease protein oppB

Query Match 100.0%; Score 32; DB 2; Length 363;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
|
|
|
|
Db 110 CRGDD 114

RESULT 5

AB0617
nicotinate phosphoribosyltransferase (EC 2.4.2.11) - Salmonella typhimurium
C:Species: Salmonella typhimurium
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: A39130
R:Vinitsky, A.; Teng, H.; Grubmeyer, C.T.
J. Bacteriol. 173, 536-540, 1991
A:Title: Cloning and nucleic acid sequence of the Salmonella typhimurium pncB gene and a
A:Reference number: A39130; MUID:91100340; PMID:1987148
A:Accession: A39130
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-400 <VIN>
A:Cross-references: UNIPROT:P22253; GB:M55986; NID:gl54268; PIDN:AAA27190.1; PID:gl54268
C:Genetics:
A:Gene: pncB
C:Function:
A:Pathway: nicotinate and nicotinamide metabolism
C:Superfamily: nicotinate phosphoribosyltransferase
C:Keywords: glycosyltransferase; pentosyltransferase

Query Match 100.0%; Score 32; DB 1; Length 400;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
|
|
|
|
Db 41 CRGDD 45

RESULT 6

JQ0756
nicotinate phosphoribosyltransferase (EC 2.4.2.11) - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: JQ0756; B64833
R:Wubbolds, M.G.; Terpstra, P.; van Beilen, J.B.; Kingma, J.; Meesters, H.A.R.; Witholt,
J. Biol. Chem. 265, 17665-17672, 1990
A:Title: Variation of cofactor levels in Escherichia coli; sequence analysis and express
A:Reference number: JQ0756; MUID:91009224; PMID:2211655
A:Accession: JQ0756
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-400 <WUB>
A:Cross-references: UNIPROT:P18133; GB:J05569; NID:gl47306; PIDN:AAA24400.1; PID:gl47307
A:Experimental source: strain GEC70
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: B64833
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-400 <BLAT>
A:Cross-references: GB:AB000195; GB:U00096; NID:gl787156; PIDN:AAC74017.1; PID:gl787162;
A:Experimental source: strain K-12, substrain MGL655
C:Genetics:
A:Gene: pncB
C:Function:

A:Pathway: nicotinate and nicotinamide metabolism
C:Superfamily: nicotinate phosphoribosyltransferase
C:Keywords: glycosyltransferase; pentosyltransferase

Query Match 100.0%; Score 32; DB 1; Length 400;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
|
|
|
|
Db 41 CRGDD 45

RESULT 7

AB0617
nicotinate phosphoribosyltransferase [imported] - Salmonella enterica subsp. enterica ser
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AB0617
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher,
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AB0617
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-400 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD05404.1; PID:gl6502165; GSPDB:GN00176
C:Genetics:
A:Gene: STY1010
C:Superfamily: nicotinate phosphoribosyltransferase

Query Match 100.0%; Score 32; DB 2; Length 400;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
|
|
|
|
Db 41 CRGDD 45

RESULT 8

F90755
nicotinate phosphoribosyltransferase [imported] - Escherichia coli (strain O157:H7, subst
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: F90755
R:Hayaishi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: F90755
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-400 <HAY>
A:Cross-references: UNIPROT:Q8XDE8; GB:BA000007; PIDN:BA834437.1; PID:gl3360473; GSPDB:GN
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECs1014
C:Superfamily: nicotinate phosphoribosyltransferase

Query Match 100.0%; Score 32; DB 2; Length 400;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
|
|
|
|
Db 41 CRGDD 45

srine/threonine-protein kinase mkl - fission yeast (Schizosaccharomyces pombe)
T38073

A; Cross-references: UNIPROT-Q92XS2; GB-AE006469; PIDN-AAK65827.1; PID-g14524331; GSPDB:G14524331.
A; Experimental source: strain 1021. megaplasmid pSymA
R/Galibert, F.; Finan, T.W.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, R.; Bataillon, J.; Chabot, P.; Cowie, L.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; Hymann, R.W.; Jones, T. L.

Science 293, 668-672, 2001

A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yen, K.
A;Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
A;Reference number: A96039; MUID:21368234; PMID:11474104

C;Genetics:
A;Content: annotation

A;Gene: Sma2147

A;Genome: plasmid

Query Match 90.6%; Score 29; DB 2; Length 296;

Best Local Similarity 80.0%; Pred. No. 1.3e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5

Db 5 CKGDD 9

RESULT 14

F96743 probable C2H2-type zinc finger protein F28P5.6 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C;Accession: F96743

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chan, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, B.; Kim, C. C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: F96743

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-413 <STO>

A;Cross-references: UNIPROT:Q9C7G8; GB:AE005173; NID:g11054410; PIDN:AAG27797.1; GSPDB:G

C;Genetics:

A;Gene: F28P5.6

A;Map position: 1

Query Match 90.6%; Score 29; DB 2; Length 413;

Best Local Similarity 80.0%; Pred. No. 1.7e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5

Db 191 CKGDD 195

RESULT 15

T16422

hypothetical protein F52C9.3 - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C;Accession: T16422

R;Favell, T.

submitted to the EMBL Data Library, November 1995

A;Description: The sequence of *C. elegans* cosmid F52C9.

A;Reference number: 218511

A;Accession: T16422

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-732 <FAV>

A;Cross-references: EMBL:U39850; NID:g1055052; PID:g1055058; PIDN:AAA81060.1; CESP:F52C9

C;Genetics:

A;Gene: CESP:F52C9.3

A;Introns: 63/1; 106/3; 148/2; 303/1; 347/1; 408/2; 650/2

Query Match 90.6%; Score 29; DB 2; Length 732;

Best Local Similarity 80.0%; Pred. No. 2.9e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5

Db 643 CKGDD 647

RESULT 16

F82078

topoisomerase IV, chain A VC2430 [imported] - *Vibrio cholerae* (strain N16961 serogroup O1

C;Species: *Vibrio cholerae*

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C;Accession: F82078

R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P. l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A;Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.

A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: F82078

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-761 <HEI>

A;Cross-references: UNIPROT:Q9KPE0; GB:AE004312; GB:AE003852; NID:g9656995; PIDN:AAF9557;

A;Experimental source: serogroup O1; strain N16961; biotype El Tor

C;Genetics:

A;Gene: VC2430

A;Map position: 1

C;Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain A; phage T4 DNA topoisomerase (I

Query Match 90.6%; Score 29; DB 2; Length 761;

Best Local Similarity 80.0%; Pred. No. 3e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5

Db 611 CKGDD 615

RESULT 17

A36453

decorsin - leech (*Macrobdella decora*)

C;Species: *Macrobdella decora*

C;Date: 08-Mar-1991 #sequence_revision 08-Mar-1991 #text_change 09-Jul-2004

C;Accession: A36453

R;Seymour J.L.; Henzel, W.J.; Nevins, B.; Stults, J.T.; Lazarus, R.A.

J. Biol. Chem. 265, 10143-10147, 1990

A;Title: decorsin. A potent glycoprotein IIB-IIIa antagonist and platelet aggregation inh

A;Reference number: A36453; MUID:90277628; PMID:2351655

A;Accession: A36453

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-39 <SEY>

A;Cross-references: UNIPROT:PI7350

Query Match 87.5%; Score 28; DB 2; Length 39;

Best Local Similarity 80.0%; Pred. No. 35;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5

Db 7 CQGDD 11

RESULT 18

DNVPE

DNA-binding protein - polyomavirus BK

N;Alternate names: agnoprotein

C;Species: Polyomavirus hominis 1 (polyomavirus BK)

C;Date: 01-Sep-1981 #sequence_revision 01-Sep-1981 #text_change 09-Jul-2004

C;Accession: B03632; B36762; JQ0627; A03632

R;Self, I.; Khoury, G.; Dhar, R.
Cell 18, 963-977, 1979
A;Title: The genome of human papovavirus BK.
A;Reference number: A03632; MUID:80090082; PMID:229976
A;Accession: B03632
A;Molecule type: DNA
A;Residues: 1-66 <SEI>
A;Cross-references: UNIPROT:P03085; GB:V01108; GB:J02038; NID:G60844; PIDN:CAA24296.1; R
A;Experimental source: strain Dunlop
R;Yang, R.C.A.; Wu, R.
Science 206, 456-462, 1979
A;Title: BK virus DNA: complete nucleotide sequence of a human tumor virus.
A;Reference number: A36762; MUID:80058557; PMID:228391
A;Accession: B36762
A;Molecule type: DNA
A;Residues: 1-66 <YAN>
A;Cross-references: GB:V01109; GB:J02039; NID:G60851; PIDN:CAA24304.1; PID:G60854
A;Experimental source: strain MM
R;Moens, U.; Sundsfjord, A.; Flegstad, T.; Traavik, T.
J. Gen. Virol. 71, 1461-1471, 1990
A;Title: BK virus early RNA transcripts in stably transformed cells: Enhanced levels in
A;Reference number: JQ0627; MUID:90324932; PMID:2165132
A;Accession: JQ0627
A;Molecule type: genomic RNA
A;Residues: 1-66 <MOE>
A;Cross-references: DDBJ:D00678; NID:G222309; PIDN:BAA00584.1; PID:G222311
A;Experimental source: strain BKT-1B
R;Jay, G.; Nomura, S.; Anderson, C.W.; Khoury, G.
Nature 291, 346-349, 1981
A;Title: Identification of the SV40 agnogene product: a DNA binding protein.
A;Reference number: A93257; MUID:81197675; PMID:6262654
A;Contents: annotation; identification of the protein
C;Comment: This protein, encoded by the agnogene, is also called agnoprotein. It may have
C;Superfamily: polyomavirus DNA-binding protein
C;Keywords: DNA binding

Query Match 87.5%; Score 28; DB 1; Length 66;
Best Local Similarity 80.0%; Pred. No. 56;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
Db 40 CRGED 44

RESULT 19
DNVPAS
DNA-binding protein - polyomavirus BK (strain AS)
N;Alternate names: agnoprotein
C;Species: Polyomavirus hominis 1 (polyomavirus BK)
A;Note: host Homo sapiens (man)
C;Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 09-Jul-2004
C;Accession: C33278
R;Tavie, J.E.; Walker, D.L.; Gardner, S.D.; Frisque, R.J.
J. Virol. 63, 901-911, 1989
A;Title: Nucleotide sequence of the human polyomavirus AS virus, an antigenic variant of
A;Reference number: A33278; MUID:89095020; PMID:2536111
A;Accession: C33278
A;Molecule type: DNA
A;Residues: 1-74 <TAV>
A;Cross-references: UNIPROT:P14998; GB:M23122; NID:G332774; PIDN:AAA46879.1; PID:G332778
C;Comment: The DNA sequence was obtained from Genbank, release 61.0.
C;Superfamily: polyomavirus DNA-binding protein
C;Keywords: DNA binding

Query Match 87.5%; Score 28; DB 1; Length 74;
Best Local Similarity 80.0%; Pred. No. 62;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
Db 48 CRGED 52

RESULT 20
I67436
interleukin-1-beta-converting enzyme and ced-3 homolog-1, long isoform - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
C;Accession: I67436
R;Flaws, J.A.; Kugu, K.; Trbovich, A.M.; DeSanti, A.; Tilly, K.I.; Hirschfield, A.N.; Tili
Endocrinology 136, 5042-5053, 1995
A;Title: Interleukin-1-beta-converting enzyme-related proteases (IRPs) and mammalian cell
nucosa cells of the ovarian follicle.
A;Reference number: I53300; MUID:96042508; PMID:7588240
A;Accession: I67436
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-182 <RES>
A;Cross-references: UNIPROT:P55215; EMBL:U34684; NID:G1004368; PIDN:AAC52260.1; PID:G1004

Query Match 87.5%; Score 28; DB 2; Length 182;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
Db 123 CRGDE 127

RESULT 21
S76342
hypothetical protein - Synechocystis sp. (strain PCC 6803)
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C;Accession: S76342
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda,
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S76342
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-234 <KAN>
A;Cross-references: UNIPROT:Q57208; EMBL:D64000; GB:AB001339; NID:G1001484; PIDN:BAA10194
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Superfamily: Mg2+-transporting ATPase

Query Match 87.5%; Score 28; DB 2; Length 234;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
Db 153 CRGDE 157

RESULT 22
B55649
TNFR-associated protein EB16 - human
C;Species: Homo sapiens (man)
C;Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 09-Jul-2004
C;Accession: B55649
R;Mosialos, G.; Birkenbach, M.; Yalamanchili, R.; VanArsdale, T.; Ware, C.; Kieff, E.
Cell 80, 389-399, 1995
A;Title: The Epstein-Barr virus transforming protein LMP1 engages signaling proteins for
A;Reference number: A55649; MUID:95163092; PMID:7859281
A;Accession: B55649
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-416 <MOS>
A;Cross-references: UNIPROT:Q13077; GB:U19261; NID:G675461; PIDN:AAA62309.1; PID:G675462

Query Match 87.5%; Score 28; DB 2; Length 416;
 Best Local Similarity 80.0%; Pred. No. 2.8e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
 |||||
 Db 57 CRGED 61

RESULT 23
 B82611
 conserved hypothetical protein XF1995 [imported] - Xylella fastidiosa (strain 9a5c)
 C:Species: Xylella fastidiosa
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
 C:Accession: B82611
 R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
 Nature 406, 151-157, 2000
 A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A:Reference number: A82515; MUID:20365717; PMID:10910347
 A:Note: for a complete list of authors see reference number A59328 below
 A:Accession: B82611
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-423 <STM>
 A:Cross-references: UNIPROT:Q9PBZ0; GB:AE004019; NID:g9107105; PIDN:AAF8479
 A:Experimental source: strain 9a5c
 R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
 as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
 submitted to GenBank, June 2000
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitejima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
 A:Authors: Martins, E.M.P.; Matsukuma, A.Y.; Menck, C.P.M.; Miracca, E.C.; Miyaki, C.Y.;
 , F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
 A:Authors: da Silva, A.C.R.; da Silva, P.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
 M.; Tshunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
 A:Reference number: A59328
 A:Contents: annotation
 C:Genetics:
 A:Gene: XF1995

Query Match 87.5%; Score 28; DB 2; Length 423;
 Best Local Similarity 80.0%; Pred. No. 2.9e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
 |||||
 Db 313 CRGDE 317

RESULT 24
 T14237
 hypothetical protein ORF427 - Synechococcus sp. (strain PCC 7002)
 C:Species: Synechococcus sp.
 A:Variety: strain PCC 7002
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
 C:Accession: T14237
 R:Klughammer, B.; Sulteneyer, D.; Badger, M.R.; Price, G.D.
 submitted to the EMBL Data Library, April 1997
 A:Description: Involvement of ndhF3, ndh3 and ORF427 genes in high affinity CO2 uptake
 A:Reference number: Z17936
 A:Accession: T14237
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-427 <KLJ>
 A:Cross-references: UNIPROT:O33751; EMBL:U97516; NID:g2232044; PID:g2232048; PIDN:AA621
 A:Experimental source: strain PCC 7002
 C:Superfamily: uncharacterized conserved protein alr4158

Query Match 87.5%; Score 28; DB 2; Length 427;
 Best Local Similarity 80.0%; Pred. No. 2.9e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
 |||||
 Db 352 CRGED 356

RESULT 25
 S77330
 hypothetical protein sll1734 - Synechocystis sp. (strain PCC 6803)
 C:Species: Synechocystis sp.
 A:Variety: PCC 6803
 C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
 C:Accession: S77330
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
 O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda,
 DNA Res. 3, 109-136, 1996
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
 S.
 A:Reference number: S74322; MUID:97061201; PMID:8905231
 A:Accession: S77330
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-431 <KAN>
 A:Cross-references: UNIPROT:P73393; EMBL:D90906; GB:AB001339; NID:g1652492; PIDN:BAAL743;
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
 C:Superfamily: uncharacterized conserved protein alr4158

Query Match 87.5%; Score 28; DB 2; Length 431;
 Best Local Similarity 80.0%; Pred. No. 2.9e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
 |||||
 Db 356 CRGED 360

RESULT 26
 A54821
 apoptosis regulator ICH-1, stimulatory form L - human
 C:Species: Homo sapiens (man)
 C:Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 09-Jul-2004
 C:Accession: A54821
 R:Wang, L.; Miura, M.; Bergeron, L.; Zhu, H.; Yuan, J.
 Cell 76, 739-750, 1994
 A:Title: Ich-1, an I κ B/c κ 3-related gene, encodes both positive and negative regulators
 A:Reference number: A54821; MUID:94373811; PMID:8087842
 A:Accession: A54821
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-435 <WAN>
 A:Cross-references: UNIPROT:P42575; GB:U13021; NID:g537291; PID:g537292
 C:Keywords: alternative splicing; apoptosis

Query Match 87.5%; Score 28; DB 2; Length 435;
 Best Local Similarity 80.0%; Pred. No. 2.9e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
 |||||
 Db 303 CRGDE 307

RESULT 27
 JC6507
 caspase-2 - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
 C:Accession: JC6507
 R:Sato, N.; Milligan, C.E.; Uchiyama, Y.; Oppenheim, R.W.
 Gene 202, 127-132, 1997
 A:Title: cloning and expression of the cDNA encoding rat caspase-2.
 A:Reference number: JC6507; MUID:98087427; PMID:9427555

A;Accession: JC6507
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-452 <SAT>
A;Cross-references: UNIPROT:O55194; GB:U77933; NID:g2769705; PIDN:AAB96379.1; PID:g27697

Query Match 87.5%; Score 28; DB 2; Length 452;
Best Local Similarity 80.0%; Pred. No. 3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
Db 320 CRGDE 324
|||||

RESULT 28
E72376
hypothetical protein TM0434 - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C;Accession: E72376; F72298
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: E72376
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-470 <ARN>
A;Cross-references: UNIPROT:Q9WVR5; GB:AE001722; NID:g4980938; PIDN:AAD3551
A;Experimental source: strain MSB8
A;Accession: F72298
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-415, 'A', 417-464, 'GS', <AR2>
A;Cross-references: GB:AE001767; GB:AE000512; NID:g4981611; PIDN:AAD36145.1; PID:g498161
C;Genetics:
A;Gene: TM0434; TM1068
C;Superfamily: melibiose-specific alpha-galactosidase

Query Match 87.5%; Score 28; DB 2; Length 470;
Best Local Similarity 80.0%; Pred. No. 3.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
Db 216 CRGDE 220
|||||

RESULT 29
T24497
hypothetical protein T05A6.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T24497
R;Thomas, K.
submitted to the EMBL Data Library, August 1995
A;Reference number: Z19899
A;Accession: T24497
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-490 <WIL>
A;Cross-references: UNIPROT:Q22200; EMBL:Z50796; PIDN:CAA90668.1; GSPDB:GN00020; CESP:T0
A;Experimental source: clone T05A6
C;Genetics:
A;Gene: CESP:T05A6.4
A;Map position: 2
A;Introns: 39/1; 125/1; 189/1; 286/3; 432/3; 474/1
C;Superfamily: Caenorhabditis elegans hypothetical protein T05A6.4

Query Match 87.5%; Score 28; DB 2; Length 452;
Best Local Similarity 80.0%; Pred. No. 3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
Db 320 CRGDE 324
|||||

RESULT 28
E72376
hypothetical protein TM0434 - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C;Accession: E72376; F72298
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: E72376
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-470 <ARN>
A;Cross-references: UNIPROT:Q9WVR5; GB:AE001722; NID:g4980938; PIDN:AAD3551
A;Experimental source: strain MSB8
A;Accession: F72298
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-415, 'A', 417-464, 'GS', <AR2>
A;Cross-references: GB:AE001767; GB:AE000512; NID:g4981611; PIDN:AAD36145.1; PID:g498161
C;Genetics:
A;Gene: TM0434; TM1068
C;Superfamily: melibiose-specific alpha-galactosidase

Query Match 87.5%; Score 28; DB 2; Length 470;
Best Local Similarity 80.0%; Pred. No. 3.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
Db 216 CRGDE 220
|||||

RESULT 29
T24497
hypothetical protein T05A6.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T24497
R;Thomas, K.
submitted to the EMBL Data Library, August 1995
A;Reference number: Z19899
A;Accession: T24497
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-490 <WIL>
A;Cross-references: UNIPROT:Q22200; EMBL:Z50796; PIDN:CAA90668.1; GSPDB:GN00020; CESP:T0
A;Experimental source: clone T05A6
C;Genetics:
A;Gene: CESP:T05A6.4
A;Map position: 2
A;Introns: 39/1; 125/1; 189/1; 286/3; 432/3; 474/1
C;Superfamily: Caenorhabditis elegans hypothetical protein T05A6.4

Query Match 87.5%; Score 28; DB 2; Length 490;
Best Local Similarity 80.0%; Pred. No. 3.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
Db 190 CRGDE 194
|||||

RESULT 30
T18896
glypican 1 precursor - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T18896; T22999
R;Lloyd, C.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19040
A;Accession: T18896
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-521 <WIL>
A;Cross-references: UNIPROT:O17900; EMBL:Z81459; PIDN:CAB03828.1; GSPDB:GN00028; CESP:F5;
A;Experimental source: clone C03H12
R;Lloyd, C.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19649
A;Accession: T22999
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-521 <WI2>
A;Cross-references: EMBL:Z81558; PIDN:CAB04542.1; GSPDB:GN00028; CESP:F59D12.4
A;Experimental source: clone F59D12
C;Genetics:
A;Gene: CESP:F59D12.4
A;Map position: X
A;Introns: 74/1; 134/3; 192/3; 247/1; 290/3; 339/3; 403/1; 450/3
C;Keywords: glycoprotein; lipoprotein; membrane protein; phosphoprotein

Query Match 87.5%; Score 28; DB 2; Length 521;
Best Local Similarity 80.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
Db 355 CRGDE 359
|||||

RESULT 31
T22137
hypothetical protein F43G6.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T22137
R;Swirnburne, J.
submitted to the EMBL Data Library, July 1995
A;Reference number: Z19522
A;Accession: T22137
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-548 <WIL>
A;Cross-references: UNIPROT:Q20367; EMBL:Z50070; NID:e1519046; PIDN:CAA90395.1; GSPDB:GN;
A;Experimental source: clone F43G6
C;Genetics:
A;Gene: CESP:F43G6.6
A;Map position: 2
A;Introns: 78/1; 120/2; 228/2; 296/3; 421/1; 547/1

Query Match 87.5%; Score 28; DB 2; Length 548;
Best Local Similarity 80.0%; Pred. No. 3.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5

```

Db          |||||:
           503 CRGDE 507

RESULT 32
T51525
hypothetical protein T20K14_80 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
C:Accession: T51525
R:Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, H.; Kotani, H.; Tabata, S.; Mew
submitted to the Protein Sequence Database, August 2000
A:Reference number: 225394
A:Accession: T51525
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-572 <SAT>
A:Cross-references: UNIPROT:Q9LFP35; EMBL:AL391143
A:Experimental source: cultivar Columbia; BAC clone T20K14
C:Genetics:
A:Map position: 5
A:Introns: 43/2; 114/1; 129/3; 181/3; 378/3; 479/3
A:Note: T20K14_80

Query Match      87.5%; Score 28; DB 2; Length 572;
Best Local Similarity 80.0%; Pred. No. 3.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CRGDD 5
      |||||:
Db      418 CRGED 422

RESULT 33
E81684
1-deoxyxylulose-5-phosphate synthase TC0608 [imported] - Chlamydia muridarum (strain Nig
C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: E81684
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: A81500; MUID:20150255; PMID:10684935
A:Accession: E81684
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-632 <TET>
A:Cross-references: UNIPROT:Q9PK62; GB:AE002329; GB:AE002160; NID:g7190640; PIDN:AAF3943
A:Experimental source: strain Nigg (MoPn)
C:Genetics:
A:Gene: TC0608
C:Superfamily: deoxyxylulose-5-phosphate synthase

Query Match      87.5%; Score 28; DB 2; Length 632;
Best Local Similarity 80.0%; Pred. No. 4.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CRGDD 5
      |||||:
Db      501 CQGGD 505

RESULT 34
T00602
hypothetical protein At2g02620 [imported] - Arabidopsis thaliana
N:Alternate names: hypothetical protein T8K22.8
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
C:Accession: T00602; G84438
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul
submitted to the EMBL Data Library, June 1998
A:Description: Arabidopsis thaliana chromosome II BAC T8K22 genomic sequence.

```

```

A:Reference number: Z14192
A:Accession: T00602
A>Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-635 <ROU>
A:Cross-references: UNIPROT:O64715; EMBL:AC004136; NID:g3184270; PID:g3184278
A:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; N
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: G84438
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-635 <STO>
A:Cross-references: GB:AE002093; NID:g3184278; PIDN:AAC18925.1; GSPDB:GN00139
C:Genetics:
A:Gene: T8K22.8; At2g02620
A:Map position: 2
A:Introns: 18/3

Query Match      87.5%; Score 28; DB 2; Length 635;
Best Local Similarity 80.0%; Pred. No. 4.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CRGDD 5
      |||||:
Db      532 CRGDE 536

RESULT 35
D84823
hypothetical protein At2g39950 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: D84823
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; N
euss, D.; Nierman, W.C.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: D84823
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-635 <STO>
A:Cross-references: UNIPROT:O04198; GB:AE002093; NID:g2088648; PIDN:AAB95280.1; GSPDB:GNC
C:Genetics:
A:Gene: At2g39950
A:Map position: 2

Query Match      87.5%; Score 28; DB 2; Length 636;
Best Local Similarity 80.0%; Pred. No. 4.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CRGDD 5
      |||||:
Db      88 CRGDE 92

RESULT 36
F71527
probable transketolase - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C:Species: Chlamydia trachomatis
C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004
C:Accession: F71527
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell,
Science 282, 754-759, 1998
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia tract
A:Reference number: A71570; MUID:9900809; PMID:9784136
A:Accession: F71527

```

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-640 <ARN>
A;Cross-references: UNIPROT:O84335; GB:AE001306; NID:G328748; PIDN:AA06752
A;Experimental source: serotype D, strain UW-3/Cx
C;Genetics:
A;Gene: dxs
C;Superfamily: deoxyxylulose-5-phosphate synthase

Query Match 87.5%; Score 28; DB 2; Length 640;
Best Local Similarity 80.0%; Pred. No. 4.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
|:|:|
Db 503 CQGDD 507

RESULT 37

T00859
Hypothetical protein At2g02690 [imported] - Arabidopsis thaliana
N;Alternate names: hypothetical protein T20F6.17; hypothetical protein T8K22.1
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C;Accession: T00859; T00595; F84439
R;Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul
submitted to the EMBL Data Library, March 1998
A;Description: Arabidopsis thaliana chromosome II BAC T20F6 genomic sequence.
A;Reference number: Z14206
A;Accession: T00859
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-657 <ROU>
A;Cross-references: UNIPROT:Q858R7; EMBL:AC002521; NID:G2947056; PID:G2947072
A;Experimental source: cultivar Columbia
R;Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul
submitted to the EMBL Data Library, June 1998
A;Description: Arabidopsis thaliana chromosome II BAC T8K22 genomic sequence.
A;Reference number: Z14192
A;Accession: T00595
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-613 <ROW>
A;Cross-references: EMBL:AC004136; NID:G3184270; PID:G3184271
A;Experimental source: cultivar Columbia
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: F84439

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-657 <SFO>
A;Cross-references: GB:AE002093; NID:G2947072; PIDN:AA05353.1; GSPDB:GN00139
C;Genetics:
A;Gene: T20F6.17; T8K22.1; At2g02690
A;Map position: 2
A;Introns: 471/3; 489/2; 560/3; 583/3

Query Match 87.5%; Score 28; DB 2; Length 657;
Best Local Similarity 80.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
|:|:|
Db 500 CRGDE 504

RESULT 38

E70770
probable atp-dependent helicase ding - Mycobacterium tuberculosis (strain H37RV)

C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: E70770
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: E70770
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-664 <COL>
A;Cross-references: UNIPROT:Q10640; GB:Z73902; GB:AL123456; NID:G3261576; PIDN:CAA98093.1
A;Experimental source: strain H37RV
C;Genetics:
A;Gene: ding

Query Match 87.5%; Score 28; DB 2; Length 664;
Best Local Similarity 80.0%; Pred. No. 4.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
|:|:|
Db 517 CQGDD 521

RESULT 39

JH0798
fasciclin IV precursor - American bird grasshopper
C;Species: Schistocerca americana (American bird grasshopper)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C;Accession: JH0798
R;Kolodkin, A.; Matthes, D.J.; O'Connor, T.P.; Patel, N.H.; Admon, A.; Bentley, D.; Good
Neuron 9, 831-845, 1992
A;Title: Fasciclin IV: Sequence, expression, and function during growth cone guidance in
A;Reference number: JH0798; MUID:93040225; PMID:1418998
A;Accession: JH0798
A;Molecule type: mRNA
A;Residues: 1-730 <KOL>
A;Cross-references: UNIPROT:Q26473; GB:L00709; NID:G160844; PID:G160845
A;Experimental source: embryo
C;Comment: This protein plays a role in growth cone guidance in the developing central ne
C;Keywords: glycoprotein; transmembrane protein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-730/Product: fasciclin IV #status predicted <MAT>
F;23-627/Domain: extracellular #status predicted <EXT>
F;628-652/Domain: transmembrane #status predicted <TMW>
F;653-730/Domain: intracellular #status predicted <INT>
F;44,71,163,267,360,539/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 87.5%; Score 28; DB 2; Length 730;
Best Local Similarity 80.0%; Pred. No. 4.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
|:|:|
Db 655 CRGED 659

RESULT 40

C83177
probable phosphotransferase protein PA3760 [imported] - Pseudomonas aeruginosa (strain P
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: C83177
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bri
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: CB31177
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-842 <STO>
 A:Cross-references: UNIPROT:Q9HXN5; GB:AE004794; GB:AE004091; NID:g9949917; PIDN:AAG0714
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA3760

Query Match 87.5%; Score 28; DB 2; Length 842;
 Best Local Similarity 80.0%; Pred. No. 5.2e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
 Db 237 CRGED 241
 |||:|

RESULT 41

MXR31

lambda 3 protein - reovirus type 1 (strain Lang)

N:Alternate names: minor core protein

C:Species: reovirus type 1

C>Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004

C:Accession: A30121

R:Wiener, J.R.; Joklik, W.K.

Virology 169, 194-203, 1989

A:Title: The sequences of the reovirus serotype 1, 2, and 3 L1 genome segments and analysis

A:Reference number: A94390; MUID:89163254; PMID:2922925

A:Accession: A30121

A:Molecule type: genomic RNA

A:Residues: 1-1267 <NIE>

A:Cross-references: UNIPROT:P17376; GB:M24734; NID:g499863

A>Note: This sequence, which matches the sequence attributed to type 1 in Fig. 2, matches the translations in entries REOLAM3P and REO3LAM3P now differ only by the sequence correction. See also PIR:MXR33.

C:Genetics:

A:Map position: segment L1

C:Superfamily: reovirus lambda 3 protein

C:Keywords: core protein

Query Match 87.5%; Score 28; DB 1; Length 1267;

Best Local Similarity 80.0%; Pred. No. 7.5e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
 Db 731 CQGGD 735
 |||:|

RESULT 42

MXR32

lambda 3 protein - reovirus type 2 (strain D5/Jones)

N:Alternate names: minor core protein

C:Species: reovirus type 2

C>Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004

C:Accession: B30121

R:Wiener, J.R.; Joklik, W.K.

Virology 169, 194-203, 1989

A:Title: The sequences of the reovirus serotype 1, 2, and 3 L1 genome segments and analysis

A:Reference number: A94390; MUID:89163254; PMID:2922925

A:Accession: B30121

A:Molecule type: genomic RNA

A:Residues: 1-1267 <NIE>

A:Cross-references: UNIPROT:P17377; GB:M31057; NID:g499865; PIDN:AAA47245.1; PID:g499866

C:Genetics:

A:Map position: segment L1

C:Superfamily: reovirus lambda 3 protein

C:Keywords: core protein

Query Match 87.5%; Score 28; DB 1; Length 1267;

Best Local Similarity 80.0%; Pred. No. 7.5e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
 Db 237 CRGED 241
 |||:|

RESULT 43

MXR33

lambda 3 protein - reovirus type 3 (strain Dearing)

N:Alternate names: minor core protein

C:Species: reovirus type 3

C>Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004

C:Accession: C30121

R:Wiener, J.R.; Joklik, W.K.

Virology 169, 194-203, 1989

A:Title: The sequences of the reovirus serotype 1, 2, and 3 L1 genome segments and analysis

A:Reference number: A94390; MUID:89163254; PMID:2922925

A:Accession: C30121

A:Molecule type: genomic RNA

A:Residues: 1-1267 <NIE>

A:Cross-references: UNIPROT:P17378; GB:M31058; NID:g499867; GB:M24734; NID:g499863; PIDN:

A>Note: This sequence, which matches the sequence attributed to type 3 in Fig. 2, matches the translations in entries REOLAM3P and REO3LAM3P now differ only by the sequence correction apparently made

C:Comment: See also PIR:MXR31.

C:Genetics:

A:Map position: segment L1

C:Superfamily: reovirus lambda 3 protein

C:Keywords: core protein

Query Match 87.5%; Score 28; DB 1; Length 1267;

Best Local Similarity 80.0%; Pred. No. 7.5e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
 Db 731 CQGGD 735
 |||:|

RESULT 44

T37533

coronin-like protein - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C:Accession: T37533

R:Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.

submitted to the EMBL Data Library, August 1997

A:Reference number: Z21721

A:Accession: T37533

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-1323 <MUR>

A:Cross-references: UNIPROT:O13686; EMBL:Z98595; PIDN:CAB11184.1; GSPDB:GNO0066; SPDB:SP1

A:Experimental source: strain 972h-; cosmid c11E3

C:Genetics:

A:Gene: SPDB:SPAC11E3.05

A:Map position: 1

RESULT 45

T23298

hypothetical protein K04B12.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T23298

Query Match 87.5%; Score 28; DB 2; Length 1323;

Best Local Similarity 80.0%; Pred. No. 7.8e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
 Db 514 CRGED 518
 |||:|

R;Steward, C.
submitted to the EMBL Data Library, December 1996
A;Reference number: Z19723
A;Accession: T23298
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1806 <WIL>
A;Cross-references: UNIPROT:O45657; EMBL:Z83232; PIDN:CA805755.1; GSPDB:GN00020; CESP:K04B12
A;Experimental source: clone K04B12
C;Genetics:
A;Gene: CESP:K04B12.1
A;Map position: 2
A;Introns: 24/2; 77/3; 294/2; 378/3; 433/3; 478/3; 523/3; 743/1; 794/3; 1049/3; 1414/3;

Query Match 87.5%; Score 28; DB 2; Length 1806;

Best Local Similarity 80.0%; Pred. No. 1e+03; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5

Db 305 CRGED 309

RESULT 46

T18524
scavenger receptor cysteine-rich protein homolog srcrm2 - Geodia cydonium
C;Species: Geodia cydonium
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T18524
R;Blumbach, B.; Pancer, Z.; Diehl-Seifert, B.; Steffen, R.; Muenkner, J.; Mueller, I.; M
J. Cell Sci. 111, 2635-2644, 1998
A;Title: The putative sponge aggregation receptor: Isolation and characterisation of a m
A;Reference number: Z18947; MUID:98369060; PMID:9701562
A;Accession: T18524
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-2043 <BLU>
A;Cross-references: UNIPROT:O96943; EMBL:Y14953; NID:e1364818; PID:e1364819; PIDN:CAA751
C;Genetics:
A;Gene: SRCRM2

Query Match 87.5%; Score 28; DB 2; Length 2043;

Best Local Similarity 80.0%; Pred. No. 1.1e+03; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5

Db 1440 CRGDE 1444

RESULT 47

T18397
protein CTRP - malaria parasite (Plasmodium falciparum)
C;Species: Plasmodium falciparum
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T18397
R;Trottein, F.; Trigilia, T.; Cowman, A.F.
Mol. Biochem. Parasitol. 74, 129-142, 1995
A;Title: Molecular cloning of a gene from Plasmodium falciparum that codes for a protein
A;Reference number: Z18926; MUID:96360471; PMID:8719155
A;Accession: T18397
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-2098 <PRO>
A;Cross-references: UNIPROT:Q25757; EMBL:U34363; NID:gl098897; PID:gl098898; PIDN:AAC469

Query Match 87.5%; Score 28; DB 2; Length 2098;

Best Local Similarity 80.0%; Pred. No. 1.2e+03; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5

Db 293 CQGGD 297

RESULT 48

T18995
hypothetical protein C06B8.7 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T18995
R;Steward, C.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19057
A;Accession: T18995
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-3036 <WIL>
A;Cross-references: UNIPROT:O17575; EMBL:Z81463; PIDN:CA803852.1; GSPDB:GN00023; CESP:CO6B8
A;Experimental source: clone C06B8
C;Genetics:
A;Gene: CESP:C06B8.7
A;Map position: 5
A;Introns: 102/3; 157/2; 288/2; 439/3; 542/1; 591/3; 798/3; 835/3; 1139/1; 1204/3; 1325/1

Query Match 87.5%; Score 28; DB 2; Length 3036;

Best Local Similarity 80.0%; Pred. No. 1.1e+03; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5

Db 2068 CRGDE 2072

RESULT 49

T18FBP
proteinase inhibitor - flesh fly (Sarcophaga bullata)
C;Species: Sarcophaga bullata
C;Date: 07-Feb-1992 #sequence_revision 22-Jul-1994 #text_change 09-Jul-2004
C;Accession: A37294
R;Papayannopoulos, I.A.; Biemann, K.
Protein Sci. 1, 278-288, 1992
A;Title: Amino acid sequence of a protease inhibitor isolated from Sarcophaga bullata det
A;Reference number: A37294; MUID:93284121; PMID:1304909
A;Accession: A37294
A;Molecule type: protein
A;Residues: 1-57 <PAP>
A;Cross-references: UNIPROT:P26228
A;Experimental source: hemolymph
C;Superfamily: basic proteinase inhibitor; animal Kunitz-type proteinase inhibitor homolo
C;Keywords: serine proteinase inhibitor
F;6-56/Domain: animal Kunitz-type proteinase inhibitor homology <BPI>
F;6-56,15-39,31-52/Disulfide bonds: #status predicted
F;16/Inhibitory site: Arg (chymotrypsin) #status predicted

Query Match 84.4%; Score 27; DB 1; Length 57;

Best Local Similarity 80.0%; Pred. No. 79; Mismatches 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5

Db 39 CRGND 43

RESULT 50

AI3045
hypothetical protein Atu39p1 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AI3045
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuttyavin, T.; Levy, R.; Li, M.; McClellan
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E
ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; MUID:21608550; PMID:11743193

A:Accession: AF3045

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-130 <KUR>

A:Cross-references: UNIPROT:Q8UBV7; GB:AE008689; PIDN:AAL44783.1; PID:gl7742422; GSPDB:C

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: Atu3981

A:Map position: linear chromosome

Query Match 84.4%; Score 27; DB 2; Length 130;

Best Local Similarity 80.0%; Pred. No. 1.6e+02;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CRGDD 5

Db 60 CEGDD 64

RESULT 51

HmrR heavy metal dependent transcription regulator [imported] - Sinorhizobium meliloti

C:Species: Sinorhizobium meliloti

C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001

C:Accession: E95330

R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows

; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.

Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001

A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti

A:Reference number: A95262; MUID:21396509; PMID:11481432

A:Accession: E95330

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-138 <KUR>

A:Cross-references: GB:AE006469; PIDN:AAK65207.1; PID:gl4523654; GSPDB:GN00165

A:Experimental source: strain 1021, megaplasmid pSymA

R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,

pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,

hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.

A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A:Reference number: A96039; MUID:21368234; PMID:11474104

A:Contents: annotation

C:Genetics:

A:Gene: hmrR

A:Genome: plasmid

C:Superfamily: transcription repressor glnR

Query Match

84.4%; Score 27; DB 2; Length 138;

Best Local Similarity 80.0%; Pred. No. 1.7e+02;

Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CRGDD 5

Db 104 CEGDD 108

RESULT 52

C95969

probable transcription regulator, merR family protein [imported] - Sinorhizobium meliloti

C:Species: Sinorhizobium meliloti

C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004

C:Accession: C95969

R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan

Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A:Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo

A:Reference number: A95842; MUID:21396508; PMID:11481431

A:Accession: C95969

A>Status: preliminary

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

A:Gene: ynfP

A:Superfamily: transcription repressor glnR

A:Contents: annotation

C:Genetics:

C;Keywords: transmembrane protein

F;7-23/Domain: transmembrane #status predicted <TM>

Query Match 84.4%; Score 27; DB 2; Length 164;
Best Local Similarity 80.0%; Pred. No. 2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
|
|
|
|
Db 63 CRGDN 67

RESULT 55

A70734

Probable cutinase precursor - Mycobacterium tuberculosis (strain H37RV)

C;Species: Mycobacterium tuberculosis

C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000

C;Accession: A70734

R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, M.; Holroyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A;Authors: Sgares, R.; Sultston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A;Reference number: A70500; MUID:98295987; PMID:9634230

A;Accession: A70734

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-219 <COL>

A;Cross-references: GB:277163; GB:AL123456; NID:g3261610; PIDN:CAB00997.1; PID:g1449318

A;Experimental source: strain H37RV

C;Genetics:

A;Gene: Rv2301

C;Superfamily: cutinase

Query Match 84.4%; Score 27; DB 2; Length 219;
Best Local Similarity 80.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CRGDD 5
|
|
|
|
Db 174 CHGDD 178

RESULT 56

A44805

eggshell protein precursor - fluke (Schistosoma haematobium)

C;Species: Schistosoma haematobium

C;Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 09-Jul-2004

C;Accession: A44805; C44805

R;Bobek, L.A.; LoVerde, P.T.; Rekosh, D.M.

Exp. Parasitol. 68, 17-30, 1989

A;Title: Schistosoma haematobium: analysis of eggshell protein genes and their expression

A;Reference number: A44805; MUID:89137380; PMID:2917627

A;Accession: A44805

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-220 <BOB>

A;Cross-references: UNIPROT:O96853; GB:M27659

A;Status: preliminary

A;Accession: C44805

A;Molecule type: DNA

A;Residues: 1-70,96-140,142-220 <BO2>

A;Cross-references: GB:M27658; NID:g160978; PID:g160979

C;Superfamily: fluke eggshell protein

C;Keywords: egg shell

F;1-18/Domain: signal sequence #status predicted <SIG>

F;19-220/Product: eggshell protein #status predicted <MAT>

Query Match 84.4%; Score 27; DB 2; Length 220;
Best Local Similarity 80.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CRGDD 5
|
|
|
|
Db 118 CHGDD 122

RESULT 57

F64859

Putative lambdaoid prophage e14 repressor protein C2 - Escherichia coli (strain K-12) cryI

C;Species: Escherichia coli

C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004

C;Accession: F64859

R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col

.A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A;Title: The complete genome sequence of Escherichia coli K-12.

A;Reference number: A64720; MUID:97426617; PMID:9278503

A;Accession: F64859

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-224 <BLAT>

A;Cross-references: UNIPROT:P75974; GB:AE000214; GB:U00096; NID:g1787382; PIDN:AAC74229.1

A;Experimental source: strain K-12, substrain MG1655

C;Genetics:

A;Genome: cryptic prophage e14

C;Superfamily: repressor protein cI

Query Match 84.4%; Score 27; DB 2; Length 224;
Best Local Similarity 80.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CRGDD 5
|
|
|
|
Db 158 CHGDD 162

RESULT 58

T15553

hypothetical protein C18B2.2 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C;Accession: T15553

R;Stellies, L.

submitted to the EMBL Data Library, November 1995

A;Description: The sequence of C. elegans cosmid C18B2.

A;Reference number: Z18368

A;Accession: T15553

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-255 <STE>

A;Cross-references: EMBL:U40413; NID:g1065467; PID:g1065472; PIDN:AAA81402.1; CESP:C18B2.

C;Genetics:

A;Gene: CESP.C18B2.2

A;Introns: 75/3; 108/3; 157/2; 206/2

Query Match 84.4%; Score 27; DB 2; Length 255;
Best Local Similarity 80.0%; Pred. No. 3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
|
|
|
|
Db 159 CRGDN 163

RESULT 59

C56084

interleukin-beta converting enzyme delta isozyme - human

C;Species: Homo sapiens (man)

C;Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 09-Jul-2004

C;Accession: C56084

R;Alnemri, E.S.; Fernandes-Alnemri, T.; Litwack, G.

J. Biol. Chem. 270, 4312-4317, 1995

A;Title: Cloning and expression of four novel isoforms of human interleukin-beta convert

A;Reference number: A56084; MUID:95181414; PMID:7876192

A;Accession: C56084
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-263 <ALN>
A;Cross-references: UNIPROT:P29466; GB:U13699; NID:g717043; PIDN:AAC50109.1; PID:g717044
C;Genetics:
A;Gene: IL1BCE
C;Keywords: alternative splicing

Query Match 84.4%; Score 27; DB 2; Length 263;
Best Local Similarity 80.0%; Pred. No. 3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 192 CRGDN 196

RESULT 60
AF0402
transketolase (EC 2.2.1.1) [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C;Accession: AF0402
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
deno-Tarrag, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
illo, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AF0402
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-276 <KUR>
A;Cross-references: UNIPROT:Q82BT1; GB:AL590842; PIDN:CAC92546.1; PID:gl5981244; GSPDB:G
C;Genetics:
A;Gene: YPO3314
C;Superfamily: transketolase, N-terminal subunit; ferredoxin 2[4Fe-4S] homology
C;Keywords: transferase

Query Match 84.4%; Score 27; DB 2; Length 276;
Best Local Similarity 80.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
Db 213 CRGND 217

RESULT 61
F85835
hypothetical protein Z3252 [imported] - Escherichia coli (strain O157:H7, substrain EDL935)
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: F85835
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: AB5480; MUID:21074935; PMID:1206551
A;Accession: F85835
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-299 <STO>
A;Cross-references: UNIPROT:Q8X7H7; GB:AE005174; NID:gl2516292; PIDN:AGS7146.1; GSPDB:G
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z3252

Query Match 84.4%; Score 27; DB 2; Length 299;
Best Local Similarity 80.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

A;Molecule type: DNA
A;Residues: 1-323 <SPT>
A;Cross-references: UNIPROT:Q9KBI4; GB:AP001513; GB:BA000004; NID:g10174345; PIDN:BA0566
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH1943

Query Match 84.4%; Score 27; DB 2; Length 323;
Best Local Similarity 80.0%; Pred. No. 3.7e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CRGDD 5
|
|
|
Db 248 CRGDD 252

RESULT 65
GVZ2ML
T2 protein - myxoma virus (strain Lausanne)
C;Species: myxoma virus
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C;Accession: A40566
R;Upton, C.; Macen, J.L.; Schreiber, M.; Mcfadden, G.
Virolgy 184, 370-382, 1991
A;Title: Myxoma virus expresses a secreted protein with homology to the tumor necrosis factor- α receptor
A;Reference number: A40566; MUID:91335768; PMID:1651597
A;Accession: A40566
A;Molecule type: DNA
A;Residues: 1-326 <UPT>
A;Cross-references: UNIPROT:P29825; GB:M95181; GB:M37976; NID:g332309; PIDN:AAA46632.1;
C;Superfamily: TNF-alpha receptor-II, vaccinia C22L type; NGF receptor repeat homology
C;Keywords: glycoprotein
F;64-105/Domain: NGF receptor repeat homology <NG2>
F;106-147/Domain: NGF receptor repeat homology <NG3>
F;66,181,205,238/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 84.4%; Score 27; DB 1; Length 326;
Best Local Similarity 80.0%; Pred. No. 3.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
|
|
|
Db 28 CRGDD 32

RESULT 66
B42476
hypothetical protein 14.1 - Salmonella typhi (strain Ty2)
C;Species: Salmonella typhi
C;Date: 24-Jul-1992 #sequence_revision 24-Jul-1992 #text_change 09-Jul-2004
C;Accession: B42476
R;Liu, D.; Verna, N.K.; Romana, L.K.; Reeves, P.R.
J. Bacteriol. 173, 4814-4819, 1991
A;Title: Relationships among the rfb regions of Salmonella serovars A, B, and D.
A;Reference number: A42476; MUID:91310590; PMID:1856174
A;Accession: B42476
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-333 <LIU>
A;Cross-references: UNIPROT:Q99192; GB:M65054; NID:g154334; PIDN:AAB49395.1; PID:g154336

Query Match 84.4%; Score 27; DB 2; Length 333;
Best Local Similarity 80.0%; Pred. No. 3.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
|
|
|
Db 221 CRGDD 225

RESULT 67
S15309
hypothetical protein 14.1 - Salmonella typhimurium

C;Species: Salmonella typhimurium
C;Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S15309
R;Jiang, X.M.; Neal, B.; Santiago, F.; Lee, S.J.; Romana, L.K.; Reeves, P.R.
Mol. Microbiol. 5, 695-713, 1991
A;Title: Structure and sequence of the rfb (O antigen) gene cluster of Salmonella serovar
A;Reference number: S15296; MUID:91260454; PMID:1710759
A;Accession: S15309
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-333 <MOL>
A;Cross-references: UNIPROT:P26401

Query Match 84.4%; Score 27; DB 2; Length 333;
Best Local Similarity 80.0%; Pred. No. 3.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
|
|
|
Db 221 CRGDD 225

RESULT 68
AC0766
probable glycosyltransferase rfbV [imported] - Salmonella enterica subsp. enterica serovar
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Jul-2004
C;Accession: AC0766
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AC0766
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-333 <PAR>
A;Cross-references: UNIPROT:Q99192; GB:AL513382; PIDN:CAD02449.1; PID:g16503316; GSPDB:GN
A;Gene: rfbV

Query Match 84.4%; Score 27; DB 2; Length 333;
Best Local Similarity 80.0%; Pred. No. 3.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
|
|
|
Db 221 CRGDD 225

RESULT 69
H70517
probable lppD protein - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C;Accession: H70517
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: H70517
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-343 <COL>
A;Cross-references: GB:Z97193; GB:AL123456; NID:g3261816; PIDN:CAB10035.1; PID:e324870; I
A;Experimental source: strain H37RV
C;Genetics:

A:Gene: lppD

Query Match 84.4%; Score 27; DB 2; Length 343;
Best Local Similarity 80.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
|
|
|
|
Db 166 CRGDN 170

RESULT 70

H70722
Probable transferase - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: H70722
R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Ruter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A: Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A: Reference number: A70500; MUID: 98295987; PMID: 9634230
A: Accession: H70722
A: Status: preliminary; nucleic acid sequence not shown; translation not shown
A: Molecule type: DNA
A: Residues: 1-346 <COL>
A: Cross-references: UNIPROT: Q50587; GB: A123456; NID: g3261623; PIDN: CAB01396.
A: Experimental source: strain H37Rv
C: Geneticks:
A: Gene: RV1520

Query Match 84.4%; Score 27; DB 2; Length 346;
Best Local Similarity 80.0%; Pred. No. 3.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CRGDD 5
|
|
|
|
Db 88 CRGDD 92

RESULT 71

F84550
Hypothetical protein At2g17310 [imported] - Arabidopsis thaliana
C: Species: Arabidopsis thaliana (mouse-ear cress)
C: Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C: Accession: F84550
R: Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, A.
Nature 402, 761-768, 1999
A: Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A: Reference number: A84420; MUID: 20083487; PMID: 10617197
A: Accession: F84550
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-370 <STO>
A: Cross-references: UNIPROT: Q8LL17; GB: AE002093; NID: g2262162; PIDN: AAB86508.1; GSPDB: GN
C: Geneticks:
A: Gene: At2g17310
A: Map position: 2
C: Superfamily: Arabidopsis thaliana hypothetical protein F24M12.210

Query Match 84.4%; Score 27; DB 2; Length 370;
Best Local Similarity 80.0%; Pred. No. 4.1e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CRGDD 5
|
|
|
|
Db 322 CNGDD 326

RESULT 72
AB4972

Nicotinate phosphoribosyltransferase (EC 2.4.2.11) [imported] - Buchnera sp. (strain APS)
C: Species: Buchnera sp.
C: Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C: Accession: AB4972
R: Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
Nature 407, 81-86, 2000
A: Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. At
A: Reference number: AB4930; MUID: 20445173; PMID: 10993077
A: Accession: AB4972
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-399 <STO>
A: Cross-references: GB: AP000398; GSPDB: GN00144
A: Experimental source: strain APS
C: Geneticks:
A: Gene: pncB; BU361
C: Superfamily: nicotinate phosphoribosyltransferase
C: Keywords: glycosyltransferase; pentosyltransferase

Query Match 84.4%; Score 27; DB 2; Length 399;
Best Local Similarity 80.0%; Pred. No. 4.4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
|
|
|
|
Db 41 CRGDN 45

RESULT 73

C64398
Hypothetical protein homolog MJ0787 - Methanococcus jannaschii
C: Species: Methanococcus jannaschii
C: Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C: Accession: C64398
R: Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, C.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A: Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
A: Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii.
A: Reference number: A64300; MUID: 96337999; PMID: 8688087
A: Accession: C64398
A: Status: preliminary; nucleic acid sequence not shown; translation not shown
A: Molecule type: DNA
A: Residues: 1-504 <BUL>
A: Cross-references: UNIPROT: Q58197; GB: U67523; GB: L77117; NID: g2826319; PIDN: AAB98783.1;
C: Geneticks:
A: Map position: FOR710775-712289

Query Match 84.4%; Score 27; DB 2; Length 504;
Best Local Similarity 80.0%; Pred. No. 5.4e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CRGDD 5
|
|
|
|
Db 465 CNGDD 469

RESULT 74

AD3382
ABC transporter-associated protein BMEI1042 [imported] - Brucella melitensis (strain 16M)
C: Species: Brucella melitensis
C: Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C: Accession: AD3382
R: DelVecchio, V.G.; Kapural, V.; Redkar, R.J.; Patra, G.; Muler, C.; Los, T.; Ivanova, I.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letessier, Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A: Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A: Reference number: AD3252; PMID: 11756688
A: Accession: AD3382
A: Status: preliminary

A:Molecule type: DNA
A:Residues: 1-507 <KUR>
A:Cross-references: UNIPROT:Q8YGW3; GB:AE008917; PIDN:AAL52223.1; PID:gl7983006; GSPDB:C
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME11042
A:Map position: 1
C:Superfamily: conserved hypothetical protein b1683

Query Match 84.4%; Score 27; DB 2; Length 507;
Best Local Similarity 80.0%; Pred. No. 5.4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
|||
Db 319 CRGDN 323

RESULT 75
E27793
methyl coenzyme M reductase (EC 1.8.-.-) alpha chain - Methanococcus vannielii
C:Species: Methanococcus vannielii
C>Date: 19-May-1989 #sequence_revision 19-May-1989 #text_change 09-Jul-2004
C:Accession: E27793
R:Cram, D.S.; Sherf, B.A.; Libby, R.T.; Mattaliano, R.J.; Ramachandran, K.L.; Reeve, J.N.
Proc. Natl. Acad. Sci. U.S.A. 84, 3992-3996, 1987
A:Title: Structure and expression of the genes, mcrBDCGA, which encode the subunits of c
A:Reference number: A94157
A:Accession: E27793
A:Molecule type: DNA
A:Residues: 1-553 <CRA>
A:Cross-references: UNIPROT:P07961
A>Note: the authors translated the codon GAA for residue 396 as Asp
C:Genetics:
A:Gene: mcrA
C:Superfamily: methyl coenzyme M reductase alpha chain
C:Keywords: metalloprotein; methanogenesis; oxidoreductase
F:151/Binding site: coenzyme F430 nickel (gln) (axial ligand) #status predicted
F:336/Active site: Tyr #status predicted
F:447/Binding site: coenzyme M (Tyr) #status predicted
F:484/Binding site: coenzyme B (Asn) #status predicted

Query Match 84.4%; Score 27; DB 2; Length 553;
Best Local Similarity 80.0%; Pred. No. 5.8e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CRGDD 5
|||
Db 85 CRGDD 89

Search completed: September 7, 2005, 20:10:10
Job time : 54 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2005, 19:55:25 ; Search time 167 Seconds
(without alignments)
15.332 Million cell updates/sec

Title: US-10-812-238B-41
Perfect score: 32
Sequence: 1 CRGDD 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : Uniprot_03: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	32	100.0	140	2	Q822Y9
2	32	100.0	174	2	Q6ZHJ9
3	32	100.0	187	2	Q9LECI
4	32	100.0	199	2	Q45047
5	32	100.0	215	2	Q9MCC2
6	32	100.0	220	2	Q9LGG1
7	32	100.0	230	2	Q6D9R3
8	32	100.0	255	1	Y166 METJA
9	32	100.0	280	2	Q9ZMQ4
10	32	100.0	289	2	Q9X6W2
11	32	100.0	304	1	HEM3_XANAC
12	32	100.0	304	1	HEM3_XANCP
13	32	100.0	311	1	LPNS11
14	32	100.0	321	2	Q6NSI1
15	32	100.0	357	2	Q8ZOU7
16	32	100.0	357	2	Q8ZJW2
17	32	100.0	363	2	Q92VE0
18	32	100.0	372	2	Q98IE7
19	32	100.0	390	2	Q8S9A8
20	32	100.0	393	2	Q8LMH4
21	32	100.0	399	1	PNCB_ECO57
22	32	100.0	399	1	PNCB_ECOLI
23	32	100.0	399	1	PNCB_SALTI
24	32	100.0	399	1	PNCB_SALTY
25	32	100.0	400	2	Q7UD27
26	32	100.0	416	2	Q83LN3
27	32	100.0	416	2	Q8FJ98
28	32	100.0	428	2	Q7XF76
29	32	100.0	430	1	SYS_SYNY3
30	32	100.0	442	1	Q804X1
31	32	100.0	446	2	Q8S6K3

32	100.0	446	2	Q8H4V4
33	100.0	448	2	Q8VUB7
34	100.0	469	2	Q6Z4U6
35	100.0	491	2	Q7PDD6
36	100.0	497	2	Q859H6
37	100.0	497	2	Q9MCC5
38	100.0	497	2	Q9KGQ6
39	100.0	640	2	Q8QS61
40	100.0	776	2	Q73RY0
41	100.0	934	2	Q8SZX4
42	100.0	1116	1	MKH1_SCHPO
43	100.0	1145	2	Q8JKT0
44	100.0	1504	2	Q7PSM9
45	100.0	1837	2	Q9VCZ9
46	90.6	73	2	Q7QP19
47	90.6	112	2	Q69T89
48	90.6	126	2	Q726G7
49	90.6	168	2	Q8XNS3
50	90.6	225	2	Q8T399
51	90.6	264	2	Q9NSJ7
52	90.6	276	2	Q8CUR7
53	90.6	296	2	Q92XS2
54	90.6	302	2	Q8BJG3
55	90.6	319	2	Q7MX77
56	90.6	324	2	Q940P2
57	90.6	360	2	Q9LKY3
58	90.6	360	2	Q9LKY4
59	90.6	366	2	Q72MX9
60	90.6	366	2	Q8F8R3
61	90.6	386	2	Q6FIV5
62	90.6	403	2	Q6EV52
63	90.6	412	2	Q84MZ4
64	90.6	413	2	Q9C7G8
65	90.6	419	2	Q6MK81
66	90.6	434	2	Q7YZS7
67	90.6	453	2	Q6MZV4
68	90.6	519	2	Q8CEG8
69	90.6	523	2	Q6GN16
70	90.6	525	2	Q66JV8
71	90.6	525	2	Q9JIG5
72	90.6	540	2	Q8BSW2
73	90.6	549	2	Q98UH5
74	90.6	553	2	Q69KU3
75	90.6	572	2	Q6DCU1
76	90.6	593	1	UB22_HUMAN
77	90.6	633	2	Q88KQ2
78	90.6	672	2	Q8A4E1
79	90.6	737	2	Q654P9
80	90.6	761	2	Q9KPE0
81	90.6	770	2	Q654V4
82	90.6	813	2	Q8LLI3
83	90.6	834	2	Q6K2T2
84	90.6	839	2	Q9LIE5
85	90.6	1086	2	Q76C42
86	90.6	1087	2	Q8U2R8
87	90.6	1089	2	Q76C30
88	90.6	1089	2	Q76C31
89	90.6	1089	2	Q76C32
90	90.6	1089	2	Q76C33
91	90.6	1089	2	Q76C34
92	90.6	1089	2	Q76C35
93	90.6	1089	2	Q76C37
94	90.6	1089	2	Q76C38
95	90.6	1089	2	Q76C44
96	90.6	1093	2	Q8J2S0
97	90.6	1094	2	Q8J2R9
98	90.6	1097	2	Q8J2S1
99	90.6	1696	1	PKCS_BRACL
100	87.5	39	1	DECO_MACDE
101	87.5	62	2	Q6GMH9
102	87.5	66	1	DNBI_POVBK
103	87.5	66	2	Q65614
104	87.5	66	2	Q6GWN3

Q8H4V4	oryza sativ
Q8VUB7	bradyrhizob
Q6Z4U6	oryza sativ
Q7PDD6	anopheles g
Q859H6	bacterioph
Q9MCC5	bacterioph
Q9KGQ6	vibrio para
Q8QS61	pongie her
Q73RY0	mycobacteri
Q8SZX4	drosophila
Q10407	schizosacch
Q8JKT0	heliothis z
Q7PSM9	anopheles g
Q9VCZ9	drosophila
Q7QP19	giardia lam
Q69T89	oryza sativ
Q726G7	desulfovibr
Q8XNS3	clostridium
Q8T399	ciona intes
Q9NSJ7	homo sapien
Q8CUR7	oceanobacil
Q92XS2	rhizobium m
Q8BJG3	mus musculu
Q7MX77	porphyromon
Q940P2	arabidopsis
Q9LKY3	glycine max
Q9LKY4	glycine max
Q72MX9	leptospi
Q8F8R3	leptospi
Q6FIV5	candida gla
Q6EV52	ciona savi
Q84MZ4	arabidopsis
Q9C7G8	arabidopsis
Q6MK81	bdellovibri
Q7YZS7	ciona intes
Q6MZV4	homo sapien
Q8CEG8	mus musculu
Q6GN16	xenopus lae
Q66JV8	mus musculu
Q9JIG5	mus musculu
Q8BSW2	mus musculu
Q98UH5	cyprinus ca
Q69KU3	oryza sativ
Q6DCU1	xenopus lae
Q9UPT9	homo sapien
Q88KQ2	pseudomonas
Q8A4E1	bacteroides
Q654P9	oryza sativ
Q9KPE0	vibrio chol
Q654V4	oryza sativ
Q8LLI3	oryza sativ
Q6K2T2	oryza sativ
Q9LIE5	arabidopsis
Q76C42	arthrodema
Q8U2R8	microsporid
Q76C30	arthrodema
Q76C31	arthrodema
Q76C32	arthrodema
Q76C33	trichophyto
Q76C34	trichophyto
Q76C35	arthrodema
Q76C37	trichophyto
Q76C38	arthrodema
Q76C44	arthrodema
Q8J2S0	trichophyto
Q8J2R9	trichophyto
Q8J2S1	trichophyto
PKCS_BRACL	branchiosto
DECO_MACDE	macrobodella
Q6GMH9	polyomaviru
DNBI_POVBK	polyomaviru
Q65614	polyomaviru
Q6GWN3	polyomaviru

105	28	87.5	66	2	Q6GWQ7	Q6wgq7 polyomavirus	178	28	87.5	289	2	Q6R7B9	Q6r7b9 ostreid her
106	28	87.5	66	2	Q89856	Q89856 polyomavirus	179	28	87.5	298	2	Q6MJY9	Q6mly9 bdellovibri
107	28	87.5	74	1	DNBI_POVBA	P14998 polyomavirus	180	28	87.5	301	2	Q87S43	Q87s43 vibrio para
108	28	87.5	82	2	Q8MVZ2	Q8mvz2 ornithodora	181	28	87.5	307	2	Q6ZLA2	Q6zla2 oryza sativ
109	28	87.5	82	2	Q8MVZ2	Q8mvz2 ornithodora	182	28	87.5	309	2	Q7IT76	Q7it76 bacterioph
110	28	87.5	98	2	Q7XHM2	Q7xhm2 oryza sativ	183	28	87.5	309	2	Q6INT6	Q6int6 xenopus lae
111	28	87.5	105	2	Q9RJB5	Q9rjb5 streptomyce	184	28	87.5	312	1	ICE2_RAT	P55215 rattus norv
112	28	87.5	106	2	Q6KXP7	Q6kxp7 oryza sativ	185	28	87.5	312	1	LPP3_MOUSE	P951y8 mus musculu
113	28	87.5	121	2	Q8V163	Q8v163 reovirus sp	186	28	87.5	312	1	LPP3_RAT	P97544 rattus norv
114	28	87.5	121	2	Q8V164	Q8v164 reovirus sp	187	28	87.5	312	2	Q6IMX4	Q6imx4 rattus norv
115	28	87.5	121	2	Q8V166	Q8v166 reovirus sp	188	28	87.5	325	2	Q7F167	Q7f167 oryza sativ
116	28	87.5	121	2	Q8V166	Q8v166 reovirus sp	189	28	87.5	336	2	Q64V52	Q64v52 bacteroides
117	28	87.5	121	2	Q8V167	Q8v167 reovirus sp	190	28	87.5	358	2	Q7Q8Q0	Q7q8q0 anophelies g
118	28	87.5	121	2	Q8V168	Q8v168 reovirus sp	191	28	87.5	362	1	CN29_HUMAN	Q725m8 homo sapien
119	28	87.5	121	2	Q8V169	Q8v169 reovirus sp	192	28	87.5	363	2	Q8TU00	Q8tuc0 methanosarc
120	28	87.5	121	2	Q8V170	Q8v170 reovirus sp	193	28	87.5	372	2	Q88375	Q88375 measles vir
121	28	87.5	121	2	Q8V171	Q8v171 reovirus sp	194	28	87.5	389	2	Q8HIN9	Q8hin9 thermoplas
122	28	87.5	121	2	Q8V172	Q8v172 reovirus sp	195	28	87.5	390	2	Q8N1I7	Q8nli7 homo sapien
123	28	87.5	121	2	Q8V173	Q8v173 reovirus sp	196	28	87.5	399	2	Q7PYN9	Q7py9 anophelies g
124	28	87.5	121	2	Q8V174	Q8v174 reovirus sp	197	28	87.5	401	2	Q6D454	Q6d454 erwinia car
125	28	87.5	121	2	Q8V175	Q8v175 reovirus sp	198	28	87.5	404	2	Q7N621	Q7n621 photorhabdu
126	28	87.5	121	2	Q8V176	Q8v176 reovirus sp	199	28	87.5	404	2	Q6MB08	Q6mb08 parachlamyd
127	28	87.5	121	2	Q8V177	Q8v177 reovirus sp	200	28	87.5	414	2	Q87D72	Q87d72 xyella fas
128	28	87.5	121	2	Q8V178	Q8v178 reovirus sp	201	28	87.5	415	2	P97018	P97018 aeromonas j
129	28	87.5	121	2	Q8V179	Q8v179 reovirus sp	202	28	87.5	416	1	TRA1_HUMAN	Q13077 homo sapien
130	28	87.5	121	2	Q8V180	Q8v180 reovirus sp	203	28	87.5	423	2	Q9PBZ0	Q9pbz0 xyella fas
131	28	87.5	121	2	Q8V181	Q8v181 reovirus sp	204	28	87.5	424	1	ICE2_CHICK	Q88943 gallus gall
132	28	87.5	121	2	Q8V182	Q8v182 reovirus sp	205	28	87.5	427	2	O33751	O33751 synecococc
133	28	87.5	121	2	Q8V183	Q8v183 reovirus sp	206	28	87.5	430	2	Q8VPV7	Q8vpv7 synecococc
134	28	87.5	121	2	Q8V184	Q8v184 reovirus sp	207	28	87.5	431	2	P73393	P73393 synecocyst
135	28	87.5	121	2	Q8V185	Q8v185 reovirus sp	208	28	87.5	435	1	ICE2_HUMAN	P42575 homo sapien
136	28	87.5	121	2	Q8V186	Q8v186 reovirus sp	209	28	87.5	435	1	ICE2_MOUSE	P29594 mus musculu
137	28	87.5	121	2	Q8V187	Q8v187 reovirus sp	210	28	87.5	435	2	Q8JTB2	Q8jtb2 golden ide
138	28	87.5	121	2	Q8V188	Q8v188 reovirus sp	211	28	87.5	436	2	Q7UY00	Q7uy00 rhodopirell
139	28	87.5	121	2	Q8V189	Q8v189 reovirus sp	212	28	87.5	437	2	Q8DKF3	Q8dkf3 synecococc
140	28	87.5	121	2	Q8V190	Q8v190 reovirus sp	213	28	87.5	450	2	Q9XYF4	Q9xyf4 drosophila
141	28	87.5	121	2	Q8V191	Q8v191 reovirus sp	214	28	87.5	451	2	Q9FB34	Q9fb34 streptomyce
142	28	87.5	121	2	Q8V192	Q8v192 reovirus sp	215	28	87.5	452	2	Q9BUP7	Q9bup7 homo sapien
143	28	87.5	121	2	Q8V193	Q8v193 reovirus sp	216	28	87.5	452	2	O55194	O55194 rattus norv
144	28	87.5	121	2	Q8V194	Q8v194 reovirus sp	217	28	87.5	452	2	Q8C9H7	Q8c9h7 mus musculu
145	28	87.5	121	2	Q8V195	Q8v195 reovirus sp	218	28	87.5	452	2	Q8K241	Q8k241 mus musculu
146	28	87.5	121	2	Q8V196	Q8v196 reovirus sp	219	28	87.5	456	2	Q6GURS	Q6gurs thermoasus
147	28	87.5	121	2	Q8V197	Q8v197 reovirus sp	220	28	87.5	456	2	Q8S0Y3	Q8s0y3 oryza sativ
148	28	87.5	121	2	Q8V198	Q8v198 reovirus sp	221	28	87.5	457	2	Q8TGT3	Q8tg37 thermoasus
149	28	87.5	121	2	Q8V199	Q8v199 reovirus sp	222	28	87.5	466	2	Q9S5X4	Q9s5x4 thermotoga
150	28	87.5	121	2	Q8V1A0	Q8v1a0 reovirus sp	223	28	87.5	469	1	SA50_HUMAN	Q9y512 homo sapien
151	28	87.5	121	2	Q8V1A1	Q8v1a1 reovirus sp	224	28	87.5	470	2	Q9WYR5	Q9wyr5 thermotoga
152	28	87.5	121	2	Q8V1A2	Q8v1a2 reovirus sp	225	28	87.5	475	2	Q95PW2	Q95pw2 caenorhabdi
153	28	87.5	121	2	Q8V1A3	Q8v1a3 reovirus sp	226	28	87.5	480	2	Q688D5	Q688d5 oryza sativ
154	28	87.5	131	2	Q96JTO	Q96jto reovirus sp	227	28	87.5	490	2	Q22200	Q22200 caenorhabdi
155	28	87.5	131	2	Q96CB2	Q96cb2 homo sapien	228	28	87.5	495	2	Q64ZB9	Q64zb9 bacteroides
156	28	87.5	136	2	Q88QA9	Q88qa9 pseudomonas	229	28	87.5	505	2	Q96I80	Q96i80 homo sapien
157	28	87.5	137	2	Q69N67	Q69n67 oryza sativ	230	28	87.5	510	2	Q9MAB8	Q9mab8 arabidopsis
158	28	87.5	151	2	Q7QRP2	Q7qrp2 giardia lam	231	28	87.5	532	2	Q8GWT1	Q8gwt1 arabidopsis
159	28	87.5	154	2	Q8BUS2	Q8bus2 mus musculu	232	28	87.5	535	2	Q9FPH36	Q9fph36 arabidopsis
160	28	87.5	160	2	Q89DG6	Q89dg6 bradyrhizob	233	28	87.5	540	2	O76771	O76771 dictyosteli
161	28	87.5	165	2	Q6A4H5	Q6a4h5 drosophila	234	28	87.5	543	2	O17900	O17900 caenorhabdi
162	28	87.5	168	2	Q8GVK1	Q8gvk1 oryza sativ	235	28	87.5	548	2	Q20367	Q20367 caenorhabdi
163	28	87.5	177	2	Q7TV51	Q7tv51 prochloroco	236	28	87.5	572	2	Q9LFP35	Q9lfp35 arabidopsis
164	28	87.5	177	2	Q8V6L9	Q8v6l9 halovirus h	237	28	87.5	575	2	Q97BA5	Q97ba5 thermoplas
165	28	87.5	177	2	Q7TDG3	Q7tdg3 halovirus h	238	28	87.5	580	2	Q8NMP7	Q8nmp7 corynebacte
166	28	87.5	185	2	Q68IK5	Q68ik5 helicobacte	239	28	87.5	584	2	Q8EJB2	Q8ejb2 shewanella
167	28	87.5	195	2	Q9VRN5	Q9vrn5 drosophila	240	28	87.5	596	2	Q8RYR5	Q8ryr5 oryza sativ
168	28	87.5	196	2	Q6EP29	Q6ep29 oryza sativ	241	28	87.5	607	2	Q8E9M8	Q8e9m8 shewanella
169	28	87.5	210	2	Q8IR71	Q8ir71 drosophila	242	28	87.5	610	2	Q9XHZ1	Q9xhz1 oryza sativ
170	28	87.5	210	2	Q70QV2	Q70qv2 human immun	243	28	87.5	613	2	Q8S8R7	Q8s8r7 arabidopsis
171	28	87.5	220	2	Q8LI15	Q8li15 oryza sativ	244	28	87.5	616	2	Q8N9M3	Q8n9m3 homo sapien
172	28	87.5	222	2	Q7XZ47	Q7xz47 griffithsia	245	28	87.5	617	2	Q9ARR6	Q9arr6 oryza sativ
173	28	87.5	232	2	Q8BNT6	Q8bnt6 mus musculu	246	28	87.5	625	2	Q6M2V5	Q6m2v5 corynebacte
174	28	87.5	232	2	Q7XXQ1	Q7xxq1 oryza sativ	247	28	87.5	632	1	DXS_CHLMU	Q9pk62 chlamydia m
175	28	87.5	234	1	Y014_SINY3	Q57208 synecocyst	248	28	87.5	635	2	O64715	O64715 arabidopsis
176	28	87.5	237	2	Q8HV9	Q8hv9 pseudomonas	249	28	87.5	636	2	O04198	O04198 arabidopsis
177	28	87.5	289	2	Q6MLP5	Q6mlp5 bdellovibri	250	28	87.5	637	2	Q8NIU1	Q8niu1 neurospora

251	28	87.5	640	1	DXS_CHLTR	O84335	chlamydia t	324	27	84.4	164	2	Q6IU23	O6iu23	escherichia
252	28	87.5	643	2	Q9FW93	Q9fw93	oryza sativ	325	27	84.4	171	2	Q9UM08	Q9um08	homo sapien
253	28	87.5	649	2	Q6NFM8	O6ntm8	corynebacte	326	27	84.4	179	2	Q7QGM8	Q7qgm8	anopheles g
254	28	87.5	661	2	Q8JTB6	O8jtb6	striped bas	327	27	84.4	181	2	Q647Q9	Q647q9	uncultured
255	28	87.5	663	2	Q73X78	Q73x78	mycobacteri	328	27	84.4	194	2	O96853	O96853	schistosoma
256	28	87.5	664	1	DING_MYCBO	P64315	mycobacteri	329	27	84.4	201	2	Q73BP3	Q73bp3	treponema d
257	28	87.5	664	1	DING_MYCTU	P64314	mycobacteri	330	27	84.4	201	2	Q9Q8T2	Q9q8t2	rabbit fibr
258	28	87.5	700	2	Q8PMT4	O8fnt4	corynebacte	331	27	84.4	206	2	Q997C2	Q997c2	myxoma viru
259	28	87.5	706	2	Q9VHK5	Q9vhu5	drosophila	332	27	84.4	214	1	RHOJ_HUMAN	Q9h4e5	homo sapien
260	28	87.5	707	2	Q7K5U3	Q7khu3	drosophila	333	27	84.4	219	2	Q62GG6	Q62gg6	oryza sativ
261	28	87.5	710	2	Q7SBR2	Q7sbr2	neurospora	334	27	84.4	221	2	Q7Q6P0	Q7q6p0	anopheles g
262	28	87.5	730	1	SM1A_SCHAM	O26473	schistocerc	335	27	84.4	224	1	YMFK_ECOLI	P75974	escherichia
263	28	87.5	842	2	Q9HXN5	Q9hxn5	pseudomonas	336	27	84.4	224	2	O8SBF6	O8sbf6	bacterioph
264	28	87.5	846	2	Q7XVU8	Q7xvu8	oryza sativ	337	27	84.4	228	2	O3S530	O3s530	rhizobium l
265	28	87.5	845	1	NETR_HUMAN	P56730	homo sapien	338	27	84.4	228	2	Q9L596	Q9l596	heliobacter
266	28	87.5	884	2	Q7XHV0	Q7xhv0	oryza sativ	339	27	84.4	230	1	CUT2_MYCBO	P63982	mycobacteri
267	28	87.5	897	2	Q93Y95	Q93y95	zea mays (m	340	27	84.4	230	1	CUT2_MYCTU	O621g7	mycobacteri
268	28	87.5	912	2	Q8L4S1	O8l4s1	oryza sativ	341	27	84.4	235	2	O62JQ7	O62jg7	burkholderi
269	28	87.5	963	1	LRP8_HUMAN	O14114	homo sapien	342	27	84.4	235	2	O63JY8	O63jy8	burkholderi
270	28	87.5	1129	2	Q7UKD4	Q7ukd4	rhodospirell	343	27	84.4	254	2	Q9U7I8	Q9u7i8	plasmodium
271	28	87.5	1240	2	Q8VA42	O8va42	chum salmon	344	27	84.4	258	2	O6GL81	O6gl81	xenopus tro
272	28	87.5	1267	1	VL3_REOVD	P17378	reovirus ty	345	27	84.4	269	2	Q62A18	Q62a18	burkholderi
273	28	87.5	1267	1	VL3_REOVJ	P17377	reovirus ty	346	27	84.4	269	2	Q63NZ8	Q63nz8	burkholderi
274	28	87.5	1267	1	VL3_REOVL	P17376	reovirus ty	347	27	84.4	270	2	O9Q8O6	Q9q8g6	myxoma viru
275	28	87.5	1267	2	O8V5E5	O8v5e5	ndelle viru	348	27	84.4	270	2	O9Q931	Q9q931	rabbit fibr
276	28	87.5	1267	2	O8S665	O8s665	reovirus ty	349	27	84.4	276	2	O8TMZ5	O8tmz5	methanosaer
277	28	87.5	1274	2	Q8JU61	O8ju61	golden shin	350	27	84.4	276	2	O66E77	O66e77	yersinia ps
278	28	87.5	1274	2	Q9E168	O9e168	grass carp	351	27	84.4	276	2	O8ZBT1	O8zbt1	yersinia pe
279	28	87.5	1274	2	Q9E3V9	O9e3v9	grass carp	352	27	84.4	278	2	O8UV28	O8uv28	xenopus lae
280	28	87.5	1280	2	O9EPX2	O9epx2	mus musculu	353	27	84.4	278	2	O6DHF1	O6dhf1	brachydanio
281	28	87.5	1323	2	O13686	O13686	schizosacch	354	27	84.4	279	2	O8OV15	O8ovi5	mus musculu
282	28	87.5	1685	2	Q9U4A2	Q9u4a2	plasmodium	355	27	84.4	280	2	O9S698	O9s698	homo sapien
283	28	87.5	1764	2	O4S657	O4s657	caenorhabdi	356	27	84.4	280	2	O9H160	O9hi60	homo sapien
284	28	87.5	1766	2	O6BC22	O6bc22	caenorhabdi	357	27	84.4	280	2	O8JG42	O8jg42	fugu rubrip
285	28	87.5	1778	2	O9BPP7	O9bpb7	antheraea v	358	27	84.4	280	2	O8JGM9	O8jgm9	fugu rubrip
286	28	87.5	1778	2	O9GUX5	O9gux5	antheraea p	359	27	84.4	281	2	O8BSU8	O8bsu8	m mus muscu
287	28	87.5	2043	2	O96943	O96943	geodia cydo	360	27	84.4	281	2	O9ESK4	O9esk4	mus musculu
288	28	87.5	2098	2	Q2S757	Q2s757	plasmodium	361	27	84.4	283	2	Q9W057	Q9w057	drosophila
289	28	87.5	2114	2	O97267	O97267	plasmodium	362	27	84.4	283	2	O56864	O56864	yersinia en
290	28	87.5	2173	2	O9U632	O9u632	drosophila	363	27	84.4	283	2	O42284	O42284	gallus gall
291	28	87.5	2177	2	Q9VL32	O9vl32	drosophila	364	27	84.4	285	2	O7MAK7	O7mak7	wollinella s
292	28	87.5	2209	2	O97324	O97324	plasmodium	365	27	84.4	286	2	O9DBD2	O9dbd2	m mus muscu
293	28	87.5	2209	2	Q9U0G6	Q9u0g6	plasmodium	366	27	84.4	287	2	P72287	P72287	rhizobium l
294	28	87.5	2265	2	O8IA53	O8ia53	plasmodium	367	27	84.4	288	2	O89FK1	O89fk1	brachyrihizob
295	28	87.5	2287	2	O8I520	O8i520	plasmodium	368	27	84.4	291	2	O8TC72	O8tc72	homo sapien
296	28	87.5	2658	2	O8T5G0	O8t5g0	plasmodium	369	27	84.4	298	2	O7OSQ0	O7osq0	limosella a
297	28	87.5	2743	2	O8IET9	O8iet9	plasmodium	370	27	84.4	299	1	YEGS_ECOLI	P76407	escherichia
298	28	87.5	3115	2	Q9VOV1	Q9vov1	drosophila	371	27	84.4	299	2	O83VS4	O83vs4	escherichia
299	28	87.5	3118	2	O17575	O17575	caenorhabdi	372	27	84.4	299	2	O7UCB0	O7ucb0	shigella fl
300	28	87.5	3954	2	O6LEV2	O6lev2	plasmodium	373	27	84.4	299	2	O83K10	O83ki0	shigella fl
301	28	87.5	5176	2	O6JHN7	O6jhn7	saccharopol	374	27	84.4	299	2	O8FFZ2	O8ffz2	escherichia
302	27	84.4	57	1	SBPI_SARBU	P26228	sarcophaga	375	27	84.4	299	2	O8X7H7	O8x7h7	escherichia
303	27	84.4	59	2	O6ETU2	O6etu2	oryza sativ	376	27	84.4	301	2	O7OSP9	O7osp9	limosella m
304	27	84.4	75	2	O9NU96	Q7nu96	chromobacte	377	27	84.4	302	2	O64T54	O64t54	bacterioides
305	27	84.4	76	2	O96858	O96858	galleria me	378	27	84.4	308	1	Y040_BPT4	P39254	bacterioides
306	27	84.4	89	2	O8FCM0	O8fcm0	escherichia	379	27	84.4	322	2	O84H43	O84h43	alcaligenes
307	27	84.4	94	2	O7PSJ2	O7pjs2	anopheles g	380	27	84.4	323	2	O9KBI4	O9kbi4	bacterioides
308	27	84.4	95	2	O6NPFQ	O6nfg9	corynebacte	381	27	84.4	323	2	O8X3S2	O8x3s2	salmonella
309	27	84.4	98	2	O853S7	O853s7	mycobacteri	382	27	84.4	324	2	O7ZX11	O7zx11	xenopus lae
310	27	84.4	112	2	Q7USG0	Q7usg0	rhodospirell	383	27	84.4	326	1	VT2_MYXVL	P29825	myxoma viru
311	27	84.4	122	2	O8T4Y2	O8t4y2	plasmodium	384	27	84.4	326	2	O8RLZ0	O8rlz0	salmonella
312	27	84.4	124	2	O7PEZ3	O7pez3	anopheles g	385	27	84.4	326	2	O8A2D3	O8a2d3	bacterioides
313	27	84.4	129	2	Q9VQU0	Q9vqu0	drosophila	386	27	84.4	333	1	RFEV_SALTY	P26401	salmonella
314	27	84.4	130	2	O8U8V7	O8u8v7	agrobacteri	387	27	84.4	333	2	O99192	O99192	salmonella
315	27	84.4	137	2	O9KYT5	O9kvt5	streptomyce	388	27	84.4	333	2	O7CAW1	O7caw1	salmonella
316	27	84.4	140	2	O8BPD2	O8bpd2	m mus muscu	389	27	84.4	337	2	O6P603	O6p603	brachydanio
317	27	84.4	141	2	Q7MUA2	Q7muu2	porphyromon	390	27	84.4	337	2	O8O4G9	O8o4g9	brachydanio
318	27	84.4	147	1	HMRI_RHIME	P58378	rhizobium m	391	27	84.4	338	2	O8QLH5	O8qlh5	mamestra co
319	27	84.4	147	1	HMR2_RHIME	P58379	rhizobium m	392	27	84.4	340	2	O7ZF21	O7zf21	desulfovibr
320	27	84.4	153	2	Q7LZC7	O7lzc7	trachemys s	393	27	84.4	341	2	O71AF3	O71af3	mamestra co
321	27	84.4	158	2	O8LRG3	O8lrg3	oryza sativ	394	27	84.4	345	2	O6K410	O6k410	oryza sativ
322	27	84.4	159	2	O8XU46	O8xu46	raletonia s	395	27	84.4	345	2	O9CST5	O9cat5	mus musculu
323	27	84.4	164	1	YBFP_ECOLI	P75737	escherichia	396	27	84.4	346	1	YF20_MYCTU	P64863	mycobacteri

397	27	84.4	346	1	YF47_MYCBO	P64864 mycobacteri	470	27	84.4	507	2	Q8QHV1	Q8qhv1 spartina mo
398	27	84.4	351	2	Q6AZT8	Q6azt8 xenopus lae	471	27	84.4	513	2	Q83869	Q83869 narcissus l
399	27	84.4	351	2	Q801F1	Q801f1 xenopus lae	472	27	84.4	513	2	Q9DIC4	Q9dic4 bromo strea
400	27	84.4	352	2	Q9WRO6	Q9wrg6 macaca mula	473	27	84.4	514	2	Q8OBV2	Q8obv2 alpinia mos
401	27	84.4	353	2	Q9J2K2	Q9j2k2 rheus monk	474	27	84.4	517	2	Q8BLK4	Q8blk4 debaryomyce
402	27	84.4	355	2	Q8GE93	Q8ge93 mycobacteri	475	27	84.4	524	2	Q7PKK6	Q7pkk6 anopheles g
403	27	84.4	356	1	YX12_CAEEL	Q18078 caenorhabdi	476	27	84.4	527	2	Q8VIX7	Q8vix7 oat necroti
404	27	84.4	358	2	Q7TZB9	Q7tzb9 mycobacteri	477	27	84.4	527	2	Q8VIX8	Q8vix8 oat necroti
405	27	84.4	359	1	Y199_MYCTU	Q07733 mycobacteri	478	27	84.4	528	2	Q6JOM7	Q6jom7 indian card
406	27	84.4	366	2	Q90163	Q90163 christoneu	479	27	84.4	528	2	Q9QD49	Q9qd49 cardamom mo
407	27	84.4	370	2	Q8LL17	Q8ll17 arabidopsis	480	27	84.4	538	2	Q8VIX9	Q8vix9 wheat strea
408	27	84.4	371	2	Q6VTS9	Q6vts9 christoneu	481	27	84.4	538	2	Q8VIX2	Q8vix2 wheat strea
409	27	84.4	388	2	Q6DED2	Q6ded2 xenopus lae	482	27	84.4	538	2	Q8VIX0	Q8vix0 wheat strea
410	27	84.4	395	2	Q8R115	Q8r115 mus musculu	483	27	84.4	539	2	Q8VIX1	Q8vix1 wheat strea
411	27	84.4	396	2	Q8BGQ0	Q8bgq0 m mus musculu	484	27	84.4	539	2	Q7T813	Q7t813 wheat strea
412	27	84.4	396	2	Q8EP90	Q8ep90 mus musculu	485	27	84.4	540	2	Q917L5	Q917l5 drosophila
413	27	84.4	397	2	Q8BG60	Q8bg60 m mus musculu	486	27	84.4	542	2	Q7T809	Q7t809 wheat strea
414	27	84.4	397	2	Q9DC00	Q9dc00 mus musculu	487	27	84.4	543	2	Q7T811	Q7t811 wheat strea
415	27	84.4	397	2	Q8K2B2	Q8k2b2 mus musculu	488	27	84.4	543	2	Q7T812	Q7t812 wheat strea
416	27	84.4	399	1	PNCB_BUCAI	P57442 buchnera ap	489	27	84.4	543	2	Q7T814	Q7t814 wheat strea
417	27	84.4	399	2	Q9QBG5	Q9qbg5 wheat yello	490	27	84.4	543	2	Q7T815	Q7t815 wheat strea
418	27	84.4	400	2	Q91X82	Q91x82 mus musculu	491	27	84.4	543	2	Q7T816	Q7t816 wheat strea
419	27	84.4	402	2	Q93YD4	Q93yd4 nicotiana t	492	27	84.4	544	2	Q63Q00	Q63qq0 burkholderi
420	27	84.4	402	2	Q93YF5	Q93yf5 nicotiana t	493	27	84.4	544	2	Q7T810	Q7t810 wheat strea
421	27	84.4	402	2	Q93YF4	Q93yf4 nicotiana t	494	27	84.4	545	2	Q7SEP2	Q7sep2 neurospora
422	27	84.4	402	2	Q93XG74	Q93xg74 nicotiana t	495	27	84.4	549	2	Q24204	Q24204 drosophila
423	27	84.4	406	2	Q93YVC1	Q93yc1 nicotiana t	496	27	84.4	553	1	MCRA_METVA	P07961 methanococc
424	27	84.4	406	2	Q6AK56	Q6ak56 desulfotale	497	27	84.4	553	2	Q6LWZ5	Q6lwz5 methanococc
425	27	84.4	407	2	Q748R7	Q748r7 geobacter s	498	27	84.4	554	2	Q9VIB0	Q9vib0 drosophila
426	27	84.4	417	2	Q41898	Q41898 rice necros	499	27	84.4	554	2	Q7TF98	Q7tf98 cucumber ve
427	27	84.4	431	2	Q8TCW7	Q8tcw7 homo sapien	500	27	84.4	555	2	Q9QBG3	Q9qbg3 wheat yello
428	27	84.4	434	2	Q8VMP8	Q8vmp8 gossypium h	501	27	84.4	557	2	Q91DH6	Q91dh6 chinese yam
429	27	84.4	436	2	Q9NE69	Q9ne69 leishmania	502	27	84.4	558	2	Q6DFH2	Q6dfh2 xenopus lae
430	27	84.4	442	2	Q89265	Q89265 wheat strea	503	27	84.4	559	1	PP02_MOUSE	O88554 mus musculu
431	27	84.4	448	2	Q91BP9	Q91bp9 oat mosaic	504	27	84.4	559	2	Q9E7C0	Q9e7c0 cucumber ve
432	27	84.4	458	2	Q9S8T2	Q9s8t2 onion yello	505	27	84.4	561	2	Q7Y0G4	Q7y0g4 oryza sativ
433	27	84.4	469	2	Q8QXP1	Q8qxp1 onion yello	506	27	84.4	562	2	Q7TH69	Q7th69 lycoriz pot
434	27	84.4	469	2	Q8QXP2	Q8qxp2 onion yello	507	27	84.4	568	2	Q6NVM1	Q6nvm1 xenopus tro
435	27	84.4	469	2	Q8QXP3	Q8qxp3 onion yello	508	27	84.4	571	2	Q9BPQ2	Q9bpq2 drosophila
436	27	84.4	469	2	Q8U220	Q8uz20 onion yello	509	27	84.4	571	2	Q62GY2	Q62gy2 burkholderi
437	27	84.4	469	2	Q8U221	Q8uz21 onion yello	510	27	84.4	572	2	Q34203	Q24203 drosophila
438	27	84.4	469	2	Q8U222	Q8uz22 onion yello	511	27	84.4	572	2	Q9VIB1	Q9vib1 drosophila
439	27	84.4	469	2	Q8U223	Q8uz23 onion yello	512	27	84.4	573	2	Q7RZQ5	Q7rzq5 neurospora
440	27	84.4	469	2	Q8U224	Q8uz24 onion yello	513	27	84.4	575	2	Q8NAG6	Q8nag6 homo sapien
441	27	84.4	469	2	Q8U225	Q8uz25 onion yello	514	27	84.4	581	2	Q65660	Q65660 barley mild
442	27	84.4	470	2	Q8UYU1	Q8uyu1 shallot yel	515	27	84.4	582	2	Q911I2	Q911i2 cardamom mo
443	27	84.4	470	2	Q7TGA8	Q7tga8 shallot yel	516	27	84.4	585	2	Q65603	Q65603 bovine immu
444	27	84.4	472	2	Q80P32	Q80p32 chinese nar	517	27	84.4	587	1	C08B_ONCMY	Q90x85 oncorhynch
445	27	84.4	472	2	Q80P33	Q80p33 chinese nar	518	27	84.4	588	2	Q70681	Q70681 sugarcane s
446	27	84.4	476	2	Q6AS69	Q6as69 desulfotale	519	27	84.4	591	2	Q6NP04	Q6np04 drosophila
447	27	84.4	488	2	Q91QR6	Q91qr6 wheat yello	520	27	84.4	591	2	Q7Q6N2	Q7q6n2 anopheles g
448	27	84.4	489	2	Q8RVB6	Q8rvb6 oryza sativ	521	27	84.4	602	1	YH02_YEAST	P38887 saccharomyc
449	27	84.4	494	2	Q92012	Q92012 wheat spind	522	27	84.4	603	1	ALS_MOUSE	P70389 mus musculu
450	27	84.4	495	2	P89683	P89683 wheat strea	523	27	84.4	603	1	ALS_RAT	P35859 rattus norv
451	27	84.4	500	2	Q6AZL3	Q6azl3 limosella a	524	27	84.4	603	2	Q70211	Q70211 rattus norv
452	27	84.4	501	2	Q6KB68	Q6kb68 sweet potat	525	27	84.4	605	1	ALS_HUMAN	P35858 homo sapien
453	27	84.4	501	2	Q6KB69	Q6kb69 sweet potat	526	27	84.4	605	1	ALS_PAPHA	O28333 papio hamad
454	27	84.4	501	2	Q6KB70	Q6kb70 sweet potat	527	27	84.4	605	2	Q8TAY0	Q8tay0 homo sapien
455	27	84.4	501	2	Q8B323	Q8b323 sweet potat	528	27	84.4	610	2	Q70SP5	Q70sp5 limosella g
456	27	84.4	501	2	Q8B324	Q8b324 sweet potat	529	27	84.4	611	2	P99885	P99885 onion yello
457	27	84.4	501	2	Q8B325	Q8b325 sweet potat	530	27	84.4	617	2	Q8UJG0	Q8jg0 measles vir
458	27	84.4	501	2	Q8B326	Q8b326 sweet potat	531	27	84.4	617	2	Q8UJ11	Q8j11 measles vir
459	27	84.4	501	2	Q8B327	Q8b327 sweet potat	532	27	84.4	617	2	Q8JSM4	Q8j4 measles vir
460	27	84.4	501	2	Q8B328	Q8b328 sweet potat	533	27	84.4	617	2	Q8JSM5	Q8j5 measles vir
461	27	84.4	501	2	Q8B329	Q8b329 sweet potat	534	27	84.4	617	2	Q8JTA8	Q8jta8 measles vir
462	27	84.4	502	2	Q9VIY9	Q9viy9 drosophila	535	27	84.4	617	2	Q8JTA9	Q8jta9 measles vir
463	27	84.4	502	2	Q8QR17	Q8qr17 spartina mo	536	27	84.4	617	2	Q9W8B0	Q9w8b0 measles vir
464	27	84.4	503	2	Q6FZW1	Q6fzw1 bartonella	537	27	84.4	617	2	Q9WIX1	Q9wix1 measles vir
465	27	84.4	503	2	Q6G3B9	Q6g3b9 bartonella	538	27	84.4	617	2	Q9WIX3	Q9wix3 measles vir
466	27	84.4	504	1	Y787_METJA	Q58197 methanococc	539	27	84.4	617	2	Q9WIX4	Q9wix4 measles vir
467	27	84.4	507	2	Q8YGM3	Q8ygm3 brucella me	540	27	84.4	617	2	Q9WIX6	Q9wix6 measles vir
468	27	84.4	507	2	Q8G100	Q8g100 brucella su	541	27	84.4	617	2	Q9WIX8	Q9wix8 measles vir
469	27	84.4	507	2	Q6MBK1	Q6mbk1 parachlamyid	542	27	84.4	617	2	Q9WIY3	Q9wiy3 measles vir

543	27	84.4	617	2	Q9WIZ0	Q9wiz0 measles vir	616	27	84.4	2258	2	O55459	O55459 barley mild
544	27	84.4	617	2	Q9WIZ7	Q9wiz7 measles vir	617	27	84.4	2258	2	O93128	O93128 barley mild
545	27	84.4	624	2	Q8EET5	Q8eet5 shewanella	618	27	84.4	2258	2	O65654	O65654 barley mild
546	27	84.4	637	2	O6CWIJ	O6cwi1 kluyveromyces	619	27	84.4	2258	2	O68VH7	O68VH7 barley mild
547	27	84.4	638	2	O91914	O91914 brachydanio	620	27	84.4	2258	2	O68VH8	O68VH8 barley mild
548	27	84.4	644	1	DXS_PROVM	O7v7d3 prochlorococcus	621	27	84.4	2258	2	O68VH9	O68VH9 barley mild
549	27	84.4	647	2	O7Q6N1	O7q6n1 anopheles g	622	27	84.4	2258	2	O9WAA3	O9WAA3 barley mild
550	27	84.4	649	2	O7RRA2	O7rra2 plasmodium	623	27	84.4	2261	2	O71BFP1	O71BFP1 barley mild
551	27	84.4	666	2	O7QB38	O7qb38 anopheles g	624	27	84.4	2340	2	O8UZ27	O8UZ27 oat mosaic
552	27	84.4	671	1	ANK6_CHICK	F51901 gallus gall	625	27	84.4	2404	2	O90733	O90733 wheat yellow
553	27	84.4	678	2	O9QB66	O9qb66 wheat yellow	626	27	84.4	2404	2	O90BG1	O90BG1 wheat yellow
554	27	84.4	680	2	O70SP4	O70sp4 limosella m	627	27	84.4	2404	2	O90BG4	O90BG4 wheat yellow
555	27	84.4	684	2	O9L2F4	O9l2f4 arabidopsis	628	27	84.4	2407	2	O9QMA1	O9QMA1 wheat yellow
556	27	84.4	686	2	O8RZ27	O8rzz7 oryza sativ	629	27	84.4	2410	1	POL1_BAYMJ	POL1206 b genome po
557	27	84.4	687	2	O9JIL0	O9jil0 mus musculu	630	27	84.4	2410	2	O9YJW3	O9YJW3 barley yell
558	27	84.4	703	2	O7Q6S1	O7q6s1 anopheles g	631	27	84.4	2411	2	O71BD5	O71BD5 barley yell
559	27	84.4	704	2	O9BPQ3	O9bpq3 drosophila	632	27	84.4	2412	1	POL1_BAYMG	POL1574 b genome po
560	27	84.4	717	2	O9VPT4	O9vpt4 drosophila	633	27	84.4	2412	2	O70WS1	O70WS1 barley yell
561	27	84.4	719	2	O87541	O87541 barley mild	634	27	84.4	2412	2	O70WS2	O70WS2 barley yell
562	27	84.4	721	2	O9PA43	O9pa43 xyliella fas	635	27	84.4	2412	2	O70WS3	O70WS3 barley yell
563	27	84.4	727	2	O83387	O83387 maciura mos	636	27	84.4	2412	2	O70WS4	O70WS4 barley yell
564	27	84.4	746	2	O9VFP00	O9vf00 drosophila	637	27	84.4	2412	2	O70WS5	O70WS5 barley yell
565	27	84.4	754	2	O9C7M3	O9c7m3 arabidopsis	638	27	84.4	2412	2	O70WS6	O70WS6 barley yell
566	27	84.4	803	2	O6JP85	O6jpb5 neodiprion	639	27	84.4	2412	2	O70WS7	O70WS7 barley yell
567	27	84.4	824	2	O6JK76	O6jkr76 neodiprion	640	27	84.4	2547	2	P90580	P90580 plasmodium
568	27	84.4	828	2	O6Z469	O6z469 oryza sativ	641	27	84.4	2772	2	O9VAV4	O9VAV4 drosophila
569	27	84.4	832	2	O9ICV1	O9icv1 wheat yell	642	27	84.4	2776	2	O869A0	O869A0 drosophila
570	27	84.4	842	2	O9SR71	O9sr71 drosophila	643	27	84.4	2894	2	O7KRX2	O7KRX2 drosophila
571	27	84.4	845	2	O8WPP8	O8wpp8 trypanosoma	644	27	84.4	2898	2	O868Z9	O868Z9 drosophila
572	27	84.4	876	2	O869X6	O869x6 dictyosteli	645	27	84.4	2931	2	O9W2C6	O9W2C6 drosophila
573	27	84.4	892	2	O6Y138	O6y138 lactuca sat	646	27	84.4	2968	2	O8MLU9	O8MLU9 drosophila
574	27	84.4	912	2	O76NT5	O76nt5 dictyosteli	647	27	84.4	3020	2	O9E160	O9E160 wheat strea
575	27	84.4	914	2	O8C726	O8c726 arabidopsis	648	27	84.4	3023	2	O6U6H0	O6U6H0 oat necroti
576	27	84.4	940	2	O81502	O81502 arabidopsis	649	27	84.4	3034	2	O8V1Y4	O8V1Y4 wheat strea
577	27	84.4	942	1	POLG_BSTVG	O65729 b genome po	650	27	84.4	3035	2	O72347	O72347 wheat strea
578	27	84.4	959	2	O18359	O18359 drosophila	651	27	84.4	3035	2	O8V1Y3	O8V1Y3 wheat strea
579	27	84.4	985	2	O01470	O01470 caenorhabdi	652	27	84.4	3035	2	O9E161	O9E161 wheat strea
580	27	84.4	996	2	O8KTE5	O8kte5 methylobact	653	27	84.4	3093	1	POLG_BSTV1	POLG730 b genome po
581	27	84.4	1011	1	UBAL_HUMAN	P41226 homo sapien	654	27	84.4	3403	2	O7T5A6	O7T5A6 onion yello
582	27	84.4	1012	2	O9BRB2	O9brb2 homo sapien	655	27	84.4	3456	2	P89201	P89201 sweet potat
583	27	84.4	1041	2	O02442	O02442 barley mild	656	27	84.4	5202	2	O6S362	O6S362 homo sapien
584	27	84.4	1076	2	O6VAN9	O6van9 trichoplusi	657	27	84.4	6739	2	O7OYB8	O7OYB8 giardia lam
585	27	84.4	1133	2	O8MR55	O8mr55 drosophila	658	27	84.4	30	2	O83DZ9	O83DZ9 coxiella bu
586	27	84.4	1163	2	O6NS64	O6ns64 mus musculu	659	27	84.4	38	2	O8GP25	O8GP25 pseudomonas
587	27	84.4	1175	2	O9VCE7	O9vce7 drosophila	660	27	84.4	43	2	O7T5E1	O7T5E1 neurospora
588	27	84.4	1180	2	O6A025	O6a025 mus musculu	661	27	84.4	45	2	O49718	O49718 arabidopsis
589	27	84.4	1182	2	O9N8H1	O9n8h1 trypanosoma	662	27	84.4	50	2	O885V4	O885V4 pseudomonas
590	27	84.4	1194	2	O9DUM1	O9dum1 barley yell	663	27	84.4	64	2	O92159	O92159 oncorhynch
591	27	84.4	1195	2	O6BK77	O6bk77 debaryomyce	664	27	84.4	57	1	YNE2_RH1ET	YNE2149 rhizobium e
592	27	84.4	1201	2	O39510	O39510 barley mild	665	27	84.4	63	2	O8VJW8	O8VJW8 mycobacteri
593	27	84.4	1284	1	NRX4_DROME	O94887 drosophila	666	27	84.4	63	2	O7TYV1	O7TYV1 mycobacteri
594	27	84.4	1284	2	O6NN26	O6nn26 drosophila	667	27	84.4	63	2	O8ETE6	O8ETE6 mus musculu
595	27	84.4	1365	2	O7REQ5	O7req5 plasmodium	668	27	84.4	64	1	V3B_IBVB	V3B241 avian infec
596	27	84.4	1366	2	O8IOH0	O8ioh0 drosophila	669	27	84.4	64	1	V3B_IBVM	V3B138 avian infec
597	27	84.4	1420	2	O9YNB1	O9ynb1 sugarcane s	670	27	84.4	64	1	V3B_IBVP3	V3B242 avian infec
598	27	84.4	1432	2	O18647	O18647 caenorhabdi	671	27	84.4	64	2	O8S0U9	O8S0U9 oryza sativ
599	27	84.4	1470	2	O89255	O89255 wheat spind	672	27	84.4	64	2	O6DV53	O6DV53 avian infec
600	27	84.4	1496	2	O9NFV5	O9nfv5 drosophila	673	27	84.4	64	2	O778R9	O778R9 avian infec
601	27	84.4	1496	2	O9VIT9	O9vit9 drosophila	674	27	84.4	64	2	O778S4	O778S4 avian infec
602	27	84.4	1607	2	O87BK2	O87ek2 xyliella fas	675	27	84.4	64	2	O9J3Q1	O9J3Q1 avian infec
603	27	84.4	1759	2	O9XTP8	O9xtp8 plasmodium	676	27	84.4	64	2	O9J4A8	O9J4A8 avian infec
604	27	84.4	1820	2	O90153	O90153 ustilago ma	677	27	84.4	65	2	O6PV93	O6PV93 infectious
605	27	84.4	1821	2	O6Y140	O6y140 lactuca sat	678	27	84.4	66	2	O9SR59	O9SR59 drosophila
606	27	84.4	1847	2	O9ZT69	O9zt69 lactuca sat	679	27	84.4	66	2	O95VA9	O95VA9 spirometra
607	27	84.4	1905	2	O9XTP6	O9xtp6 plasmodium	680	27	84.4	66	2	O6DTJ5	O6DTJ5 zea mays (m
608	27	84.4	2042	2	O25766	O25766 plasmodium	681	27	84.4	69	2	O8S8D5	O8S8D5 arabidopsis
609	27	84.4	2081	2	O7QK87	O7qk87 anopheles g	682	27	84.4	69	2	O7VI45	O7VI45 helicobacte
610	27	84.4	2117	2	O7KSW1	O7ksw1 drosophila	683	27	84.4	70	2	O8VAV0	O8VAV0 white spot
611	27	84.4	2165	2	O9VHW9	O9vhw9 drosophila	684	27	84.4	71	2	O18256	O18256 caenorhabdi
612	27	84.4	2174	2	O9GQRO	O9gqro drosophila	685	27	84.4	75	2	O82RN8	O82RN8 streptomyc
613	27	84.4	2178	2	O8I643	O8i643 plasmodium	686	27	84.4	76	1	LC38_ARATH	LC382753 arabidopsis
614	27	84.4	2212	2	O94657	O94657 plasmodium	687	27	84.4	78	2	O622D8	O622D8 oryza sativ
615	27	84.4	2258	1	POL1_BAMMN	P90245 b genome po	688	27	84.4	80	2	O8VX96	O8VX96 pinus pinas

689	26	81.2	80	2	Q91TJ0	Q91tj0 tupaiid her	762	26	81.2	143	2	Q6TUZ4	Q6tuz4 yaba monkey
690	26	81.2	81	2	Q71TU0	Q71tu1 plethodon j	763	26	81.2	143	2	Q9DHU5	Q9dhu5 yaba-like d
691	26	81.2	84	2	Q6KIT0	Q6kit0 oryza sativ	764	26	81.2	145	2	Q7Z024	Q7z024 eunapius fr
692	26	81.2	85	2	Q6YOV0	Q6yov0 human immun	765	26	81.2	145	2	Q9KQK7	Q9kqk7 vibrio chol
693	26	81.2	87	2	Q86539	Q86539 human herpe	766	26	81.2	147	1	Y237_TROWT	P67323 tropheryma
694	26	81.2	88	2	Q9HNQ6	Q9hnq6 halobacteri	767	26	81.2	147	1	Y533_TROW8	P67324 tropheryma
695	26	81.2	88	2	Q88801	Q88801 eastern equ	768	26	81.2	147	2	Q01599	Q01599 phanerocha
696	26	81.2	92	2	Q6MPQ6	Q6mpq6 bdellovibri	769	26	81.2	147	2	Q88AB6	Q88ab6 pseudomonas
697	26	81.2	93	2	Q6N8E1	Q6n8e1 rhodospseudo	770	26	81.2	148	2	Q88AB6	Q88ab6 pseudomonas
698	26	81.2	93	2	Q7W7K6	Q7w7k6 rhodospseudo	771	26	81.2	149	2	Q95W28	Q95w28 anthonomus
699	26	81.2	93	2	Q7WKZ5	Q7wkz5 bordetella	772	26	81.2	150	2	Q6TGP9	Q6tgp9 lecanora st
700	26	81.2	93	2	Q89US7	Q89us7 bradyrhizob	773	26	81.2	152	2	Q6RZ15	Q6rzi5 aspergillus
701	26	81.2	96	2	Q6Z4N8	Q6z4n8 oryza sativ	774	26	81.2	152	2	Q19349	Q19349 caenorhabdi
702	26	81.2	96	2	Q67MW1	Q67mw1 symbiobacte	775	26	81.2	152	2	Q47026	Q47026 escherichia
703	26	81.2	101	2	Q8VJ46	Q8vj46 mycobacteri	776	26	81.2	153	2	Q6IIA9	Q6iia9 drosophila
704	26	81.2	102	2	Q7TWZ1	Q7twz1 mycobacteri	777	26	81.2	153	2	Q7V188	Q7v188 prochloroco
705	26	81.2	102	2	Q9YA21	Q9ya21 aeropyrum p	778	26	81.2	157	2	Q6MWJ3	Q6mwj3 oryza sativ
706	26	81.2	102	2	Q6W3K5	Q6w3k5 alvinella p	779	26	81.2	157	2	Q6Z7A4	Q6z7a4 oryza sativ
707	26	81.2	102	2	Q9RQY8	Q9rqy8 mycobacteri	780	26	81.2	158	2	Q8S9W1	Q8s9w1 oryza sativ
708	26	81.2	102	2	Q82YG8	Q82yg8 streptomyce	781	26	81.2	158	2	Q6VZB6	Q6vzb6 canarypox v
709	26	81.2	103	1	YGU3_YEAST	P53097 saccharomyc	782	26	81.2	160	1	LY96_HUMAN	Q9v6y9 homo sapien
710	26	81.2	104	2	Q6SX61	Q6sx61 oryza sativ	783	26	81.2	160	2	Q9N9Q8	Q9n9q8 leishmania
711	26	81.2	104	2	Q84192	Q84192 newcastle d	784	26	81.2	160	2	Q6TGR2	Q6tgr2 equus cabal
712	26	81.2	104	2	Q9WJ27	Q9wj27 newcastle d	785	26	81.2	161	2	Q7LH18	Q7lh18 phanerocha
713	26	81.2	109	2	Q7VVZ3	Q7vvz3 bordetella	786	26	81.2	161	2	Q8Z3S0	Q8z3s0 salmonella
714	26	81.2	110	2	Q9DWM8	Q9dwm8 hepatitis c	787	26	81.2	161	2	Q8ZM26	Q8zm26 salmonella
715	26	81.2	112	2	Q92ZS8	Q92zs8 human immun	788	26	81.2	162	2	Q8PSU4	Q8psu4 methanosarc
716	26	81.2	113	2	Q8H6L0	Q8h6l0 brassica ra	789	26	81.2	162	2	Q67M90	Q67m90 oryza sativ
717	26	81.2	114	2	Q7ZEM7	Q7zem7 human immun	790	26	81.2	162	2	Q6S707	Q6s707 oryza sativ
718	26	81.2	115	2	Q6GXU6	Q6gxu6 human immun	791	26	81.2	162	2	Q6X211	Q6x211 bovine herp
719	26	81.2	115	2	Q6ZEH0	Q6zeh0 human immun	792	26	81.2	163	1	TPC_BRALA	P80322 branchiosto
720	26	81.2	115	2	Q7ZEH0	Q7zeh0 human immun	793	26	81.2	164	2	Q6TGO0	Q6tgo0 lecanora na
721	26	81.2	117	2	Q9LRK1	Q9lrk1 arabisdopsis	794	26	81.2	164	2	Q6TGO1	Q6tgo1 lecanora na
722	26	81.2	117	2	Q9WNX9	Q9wnx9 human immun	795	26	81.2	164	2	Q6TGO2	Q6tgo2 lecanora fl
723	26	81.2	118	2	Q7VWQ1	Q7vwq1 bordetella	796	26	81.2	164	2	Q6TGO3	Q6tgo3 dermatocarp
724	26	81.2	118	2	Q7WAP4	Q7wap4 bordetella	797	26	81.2	164	2	Q6ZVJ9	Q6zvj9 homo sapien
725	26	81.2	118	2	Q7WJU8	Q7wj8 bordetella	798	26	81.2	164	2	P90687	P90687 branchiosto
726	26	81.2	118	2	Q92517	Q92517 beet virus	799	26	81.2	164	2	P92198	P92198 branchiosto
727	26	81.2	118	2	Q9VPH2	Q9vph2 broad bean	800	26	81.2	164	2	Q75NJ6	Q75nj6 branchiosto
728	26	81.2	119	2	Q6V2D2	Q6v2d2 potato mop-	801	26	81.2	165	2	Q6VTR5	Q6vtr5 oryza sativ
729	26	81.2	119	2	Q8OQA2	Q8oqa2 potato mop-	802	26	81.2	165	2	Q75LG3	Q75lg3 oryza sativ
730	26	81.2	119	2	Q85292	Q85292 potato mop-	803	26	81.2	166	2	Q89TW2	Q89tw2 bradyrhizob
731	26	81.2	119	2	Q91V53	Q91v53 potato mop-	804	26	81.2	166	2	Q77024	Q77024 ephydratia m
732	26	81.2	121	2	Q81NN4	Q81nn4 drosophila	805	26	81.2	168	2	Q95231	Q95231 lucilia cup
733	26	81.2	121	2	Q8CKC2	Q8ckc2 versinia pe	806	26	81.2	168	2	Q9RXQ5	Q9rxq5 deinococcus
734	26	81.2	122	2	Q7OMZ0	Q7omz0 anophelies g	807	26	81.2	170	2	Q95PD2	Q95pd2 anophelies d
735	26	81.2	123	2	Q814A6	Q814a6 caenorhabdi	808	26	81.2	172	2	Q8KL61	Q8kl61 rhizobium e
736	26	81.2	125	2	Q96TV9	Q96tv9 pleurotus o	809	26	81.2	173	2	Q8L170	Q8l170 caenorhabdi
737	26	81.2	125	2	Q8LE17	Q8le17 arabisdopsis	810	26	81.2	173	2	Q6MXZ8	Q6mxz8 serratia ma
738	26	81.2	125	2	Q7D087	Q7d087 agrobacteri	811	26	81.2	174	2	Q6RZ25	Q6rz25 westerdykel
739	26	81.2	126	1	CYTL_DROME	P23779 drosophila	812	26	81.2	175	2	Q8WU36	Q8wu36 homo sapien
740	26	81.2	126	2	Q9VMU0	Q9vmu0 drosophila	813	26	81.2	175	2	Q722G1	Q722g1 homo sapien
741	26	81.2	127	1	WAP_RABIT	P09412 cryptotagus	814	26	81.2	176	2	Q8THI6	Q8thi6 methanosarc
742	26	81.2	127	2	Q6WNZ9	Q6wnz9 bdellovibri	815	26	81.2	176	2	Q6RZ19	Q6rz19 pseudoeuroti
743	26	81.2	128	2	Q96MK0	Q96mk0 homo sapien	816	26	81.2	176	2	Q9SE46	Q9se46 arabisdopsis
744	26	81.2	131	2	Q61HG0	Q61hg0 drosophila	817	26	81.2	176	2	Q9SSN3	Q9ssn3 arabisdopsis
745	26	81.2	131	2	Q90817	Q90817 human immun	818	26	81.2	177	2	Q86BA8	Q86ba8 drosophila
746	26	81.2	133	2	Q64564	Q64564 rattus norv	819	26	81.2	177	2	Q810E7	Q810e7 mus muscucu
747	26	81.2	134	2	Q63NE8	Q63ne8 burkholderi	820	26	81.2	178	2	Q95VP9	Q95vp9 leucophaea
748	26	81.2	135	2	Q01600	Q01600 phanerocha	821	26	81.2	178	2	Q8Z4F8	Q8z4f8 oryza sativ
749	26	81.2	135	2	Q69WT7	Q69wt7 oryza sativ	822	26	81.2	178	2	Q8PJH1	Q8pjh1 xanthomonas
750	26	81.2	136	2	Q9DHU8	Q9dhu8 yaba-like d	823	26	81.2	179	2	Q6ZDS3	Q6zds3 burkholderi
751	26	81.2	138	2	Q800C2	Q800c2 crotalus vi	824	26	81.2	179	2	Q8Y1M2	Q8y1m2 raietonia s
752	26	81.2	138	2	Q800C4	Q800c4 crotalus vi	825	26	81.2	179	2	Q6QQ98	Q6qq98 avipoxvirus
753	26	81.2	139	2	Q46306	Q46306 drosophila	826	26	81.2	180	2	Q7V167	Q7v167 prochloroco
754	26	81.2	139	2	Q52854	Q52854 rhizobium l	827	26	81.2	182	2	Q27267	Q27267 methanobact
755	26	81.2	140	2	Q8S1C2	Q8s1c2 oryza sativ	828	26	81.2	182	2	Q19136	Q19136 caenorhabdi
756	26	81.2	140	2	Q8KJW3	Q8kjw3 proteus vul	829	26	81.2	184	2	Q67UW2	Q67uw2 oryza sativ
757	26	81.2	140	2	Q6JLB3	Q6jlb3 franciella	830	26	81.2	184	2	Q6BZM0	Q6bzm0 debaryomyce
758	26	81.2	140	2	Q70K60	Q70k60 gordonia we	831	26	81.2	185	2	Q96K83	Q96k83 oryza sativ
759	26	81.2	141	2	Q8CDG0	Q8cdg0 mus muscucu	832	26	81.2	185	2	Q84X17	Q84x17 arabisdopsis
760	26	81.2	142	2	Q47031	Q47031 escherichia	833	26	81.2	186	2	Q6SV38	Q6sv38 metarhizium
761	26	81.2	142	2	Q8UGR3	Q8ugr3 agrobacteri	834	26	81.2	186	2	Q7QST4	Q7qst4 gnetum gnem

835	26	81.2	187	2	Q653A3	Q653a3 oryza sativ	908	26	81.2	222	2	Q8S0Q2	Q8s0q2 oryza sativ
836	26	81.2	187	2	Q9ZF59	Q9zf59 mycobacteri	909	26	81.2	222	1	Q7VGJ2	Q7vgj2 helicobacte
837	26	81.2	187	2	Q87FA9	Q87fa9 xylella fas	910	26	81.2	223	1	Y863_METTH	Q26951 methanobact
838	26	81.2	187	2	Q9PHB6	Q9phb6 xylella fas	911	26	81.2	223	2	Q6DMR2	Q6dmr2 crassostrea
839	26	81.2	188	1	KLK3_RAT	P15950 rattus norv	912	26	81.2	223	2	Q9N5A3	Q9n5a3 caenorhabdi
840	26	81.2	188	2	Q805I9	O805i9 arabidopsis	913	26	81.2	223	2	Q6ZJC8	Q6zjc8 oryza sativ
841	26	81.2	188	2	Q9LRL9	Q9lrl9 arabidopsis	914	26	81.2	223	2	Q9LRK5	Q9lrk5 arabidopsis
842	26	81.2	188	2	Q63XE8	Q63xe8 burkholderi	915	26	81.2	224	2	Q47391	Q47391 escherichia
843	26	81.2	188	2	Q7U723	Q7u723 synecococc	916	26	81.2	224	2	Q8XAJ2	Q8xaj2 escherichia
844	26	81.2	189	2	Q8HAB3	Q8hab3 salmonella	917	26	81.2	225	1	CLD8_MOUSE	Q8x260 mus musculu
845	26	81.2	189	2	Q89EC3	Q89ec3 bradyrhizob	918	26	81.2	225	1	RPE_RHOU	P51013 rhodospirill
846	26	81.2	189	2	Q8BMZ3	Q8bmz3 mus musculu	919	26	81.2	226	2	Q9HR90	Q9hr90 halobacteri
847	26	81.2	190	2	Q6CC46	Q6cc46 yarrowia li	920	26	81.2	227	2	Q00786	Q00786 emericella
848	26	81.2	190	2	Q8S8D8	Q8s8d8 arabidopsis	921	26	81.2	228	1	LPRH_MYCBO	P65317 mycobacteri
849	26	81.2	190	2	Q6VMQ2	Q6vmq2 gossypium b	922	26	81.2	228	1	LPRH_MYCTU	P65316 mycobacteri
850	26	81.2	190	2	Q7UIK2	Q7uik2 rhodopirell	923	26	81.2	228	1	NA61_ARATH	Q9m290 arabidopsis
851	26	81.2	191	2	Q6I7V5	Q6i7v5 salmirel sci	924	26	81.2	228	2	Q75QL7	Q75ql7 bacterioph
852	26	81.2	191	2	Q8JJK3	Q8jjk3 virus phich	925	26	81.2	229	2	Q729J8	Q729j8 fungal endo
853	26	81.2	191	2	Q63IC9	Q63ic9 burkholderi	926	26	81.2	229	2	Q729J9	Q729j9 fungal endo
854	26	81.2	193	2	Q92J63	Q92j63 rhizobium m	927	26	81.2	229	2	Q729K3	Q729k3 fungal endo
855	26	81.2	195	2	Q8I9P3	Q8i9p3 aurelia aur	928	26	81.2	229	2	Q6E5T1	Q6est1 fugu rubrip
856	26	81.2	196	2	Q97CB6	Q97cb6 thermoplas	929	26	81.2	230	2	Q63P62	Q63p62 burkholderi
857	26	81.2	196	2	Q23633	Q23633 caenorhabdi	930	26	81.2	231	2	Q9CJX9	Q9cix9 pasteurella
858	26	81.2	196	2	Q8S1B1	Q8s1b1 oryza sativ	931	26	81.2	232	2	Q6IKH5	Q6ikh5 drosophila
859	26	81.2	196	2	Q8L240	Q8l240 salmonella	932	26	81.2	232	2	Q8FIF8	Q8fif8 escherichia
860	26	81.2	196	2	Q7WTH7	Q7wth7 escherichia	933	26	81.2	233	2	Q9PT51	Q9pt51 agkistrodon
861	26	81.2	198	2	Q8K282	Q8k282 mus musculu	934	26	81.2	235	1	NA50_ARATH	Q9fmr3 arabidopsis
862	26	81.2	199	2	Q6CQO7	Q6cqo7 yarrowia li	935	26	81.2	236	2	Q9U7H7	Q9u7h7 plasmodium
863	26	81.2	199	2	Q6IJA8	Q6ija8 drosophila	936	26	81.2	236	2	Q6ZF66	Q6zfe6 oryza sativ
864	26	81.2	199	2	Q8FRZ4	Q8frz4 corynebacte	937	26	81.2	237	2	Q7P0S9	Q7p0s9 chromobacte
865	26	81.2	200	2	Q6E5F6	Q6e5f6 nosema locu	938	26	81.2	237	2	Q7V4N4	Q7v4n4 prochloroco
866	26	81.2	200	2	Q84J93	Q84j93 arabidopsis	939	26	81.2	237	2	Q6AME6	Q6ame6 desulfotale
867	26	81.2	200	2	Q84JX6	Q84jx6 arabidopsis	940	26	81.2	238	1	VSP1_AGKHA	P81176 agkistrodon
868	26	81.2	200	2	Q84JX5	Q84jx5 arabidopsis	941	26	81.2	238	2	Q67W67	Q67w67 oryza sativ
869	26	81.2	200	2	Q84UG5	Q84ug5 arabidopsis	942	26	81.2	238	2	Q9LRM3	Q9lrm3 arabidopsis
870	26	81.2	201	2	Q9LRSK0	Q9lrsk0 arabidopsis	943	26	81.2	238	2	Q810E6	Q810e6 mus musculu
871	26	81.2	201	2	Q810E8	Q810e8 mus musculu	944	26	81.2	239	2	Q96UV6	Q96uv6 ophiostoma
872	26	81.2	202	2	Q6J9V1	Q6j9v1 zea mays (m	945	26	81.2	239	2	Q7P0A4	Q7p0a4 anopheles g
873	26	81.2	202	2	Q6ZFH8	Q6zfh8 oryza sativ	946	26	81.2	239	2	Q7D6T1	Q7d6t1 mycobacteri
874	26	81.2	202	2	Q9LRL0	Q9lrl0 arabidopsis	947	26	81.2	239	2	Q63275	Q63275 rattus norv
875	26	81.2	202	2	Q8RK24	Q8rk24 pseudomonas	948	26	81.2	239	2	Q6T5L0	Q6t5l0 gloydus sh
876	26	81.2	203	2	Q7SHI7	Q7shi7 neurospora	949	26	81.2	239	2	Q8AVC1	Q8avc1 xenopus lae
877	26	81.2	205	1	CN7A_SCHPO	Q9nuj7 schizosacch	950	26	81.2	240	2	Q6BZV7	Q6bzv7 yarrowia li
878	26	81.2	205	1	Q8ZMU6	Q8zmu6 salmonella	951	26	81.2	240	2	Q7PYJ1	Q7pyj1 anopheles g
879	26	81.2	206	2	Q7S676	Q7s676 neurospora	952	26	81.2	240	2	Q9GV76	Q9gv76 lumbricus t
880	26	81.2	206	2	Q8LCQ1	Q8lcq1 arabidopsis	953	26	81.2	241	2	Q7ZXM1	Q7zxm1 xenopus lae
881	26	81.2	206	2	Q9LIW6	Q9liw6 streptomyce	954	26	81.2	242	2	Q6PIU7	Q6piu7 xenopus tro
882	26	81.2	207	2	Q6MWB3	Q6mwb3 oryza sativ	955	26	81.2	243	2	Q9CZY6	Q9czy6 homo sapien
883	26	81.2	207	2	Q66NL2	Q66nl2 human immun	956	26	81.2	243	2	Q16206	Q16206 caenorhabdi
884	26	81.2	208	2	Q9S214	Q9s214 arabidopsis	957	26	81.2	244	2	Q9A3W7	Q9a3w7 caulobacter
885	26	81.2	208	2	Q93515	Q93515 salmonella	958	26	81.2	244	2	Q93WH5	Q93wh5 oryza sativ
886	26	81.2	209	1	DR1F_ARATH	Q9lnr86 arabidopsis	959	26	81.2	244	2	Q7W3G3	Q7w3g3 bordetella
887	26	81.2	209	2	Q6YWN0	Q6ywn0 oryza sativ	960	26	81.2	244	2	Q7WET3	Q7wet3 bordetella
888	26	81.2	211	2	Q9ACW9	Q9acw9 oryza sativ	961	26	81.2	246	2	Q95ZC7	Q95zc7 leishmania
889	26	81.2	212	2	Q943F9	Q943f9 oryza sativ	962	26	81.2	247	2	Q52201	Q52201 pseudomonas
890	26	81.2	213	2	Q29521	Q29521 oryctolagus	963	26	81.2	247	2	Q88RE6	Q88re6 pseudomonas
891	26	81.2	213	2	Q29522	Q29522 oryctolagus	964	26	81.2	247	2	Q6H9T4	Q6h9t4 phage phi 4
892	26	81.2	213	2	Q82IM9	Q82im9 streptomyce	965	26	81.2	249	2	Q94I62	Q94i62 oryza sativ
893	26	81.2	214	1	ST14_SOLTU	Q41495 solanum tub	966	26	81.2	249	2	Q6H7Z8	Q6h7z8 oryza sativ
894	26	81.2	214	2	Q7QGM9	Q7qgm9 anopheles g	967	26	81.2	250	2	Q6CD23	Q6cd23 yarrowia li
895	26	81.2	214	2	Q7QGN0	Q7qgn0 anopheles g	968	26	81.2	251	2	Q7EP50	Q7ep50 anopheles g
896	26	81.2	214	2	Q6H3J0	Q6h3j0 oryza sativ	969	26	81.2	251	2	Q9U7H5	Q9u7h5 plasmodium
897	26	81.2	216	2	Q9FBJ8	Q9fbj8 streptomyce	970	26	81.2	252	2	Q9LRJ9	Q9lrj9 arabidopsis
898	26	81.2	217	2	Q6ZLNK8	Q6zlnk8 homo sapien	971	26	81.2	254	2	Q95UP4	Q95up4 stomoxys ca
899	26	81.2	217	2	Q9LSJ79	Q9lsj79 arabidopsis	972	26	81.2	254	2	Q941U1	Q941u1 oryza sativ
900	26	81.2	218	2	Q6ZSC6	Q6zsc6 homo sapien	973	26	81.2	254	2	Q7N741	Q7n741 photorhabdu
901	26	81.2	218	2	Q85IN3	Q85in3 oryza sativ	974	26	81.2	254	2	Q6IQY4	Q6iqy4 mus musculu
902	26	81.2	218	2	Q9LRLK4	Q9lrlk4 arabidopsis	975	26	81.2	255	2	Q9LJW2	Q9ljw2 arabidopsis
903	26	81.2	220	2	Q08332	Q08332 lycopersico	976	26	81.2	255	2	Q9SIP6	Q9sip6 arabidopsis
904	26	81.2	220	2	Q9LVZ8	Q9lvz8 arabidopsis	977	26	81.2	256	2	Q9LRK8	Q9lrk8 arabidopsis
905	26	81.2	221	2	Q8FQ54	Q8fq54 corynebacte	978	26	81.2	256	2	Q7ML06	Q7ml06 vibrio vuln
906	26	81.2	221	2	Q8FZ32	Q8fz32 arabidops	979	26	81.2	256	2	Q7U9S0	Q7u9s0 synecococc
907	26	81.2	221	2	Q9ALZ8	Q9alz8 fremyella d	980	26	81.2	257	2	Q7Q2X4	Q7q2x4 anopheles g

981 26 81.2 257 2 Q9LRK3
 982 26 81.2 257 2 Q9LRK6
 983 26 81.2 258 2 Q8S3P3
 984 26 81.2 258 2 Q6ZDF4
 985 26 81.2 258 2 Q7X8N1
 986 26 81.2 258 2 Q9LRK2
 987 26 81.2 258 2 Q9LRK2
 988 26 81.2 258 2 Q9W2I5
 989 26 81.2 258 2 Q7TY53
 990 26 81.2 259 1 CFAD_MOUSE
 991 26 81.2 259 1 MK16_SCHMA
 992 26 81.2 259 1 RNS2_ARATH
 993 26 81.2 259 2 Q9V7E5
 994 26 81.2 259 2 Q6NPE1
 995 26 81.2 259 2 Q697D7
 996 26 81.2 260 2 Q6VPU6
 997 26 81.2 260 2 Q9FNC8
 998 26 81.2 260 2 Q7SZC3
 999 26 81.2 261 2 Q7PZP7
 1000 26 81.2 261 2 Q9VXC7

ALIGNMENTS

RESULT 1
 Q822Y9 ID Q822Y9 PRELIMINARY; PRT; 140 AA.
 AC Q822Y9;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocusNames=CCA00537;
 OS Chlamydomophila caviae.
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomophila.
 OC NCBI_TaxID=83557;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GPIC;
 RX MEDLINE=22569155; PubMed=12682364; DOI=10.1093/nar/gkg321;
 RA Read T.D., Myers G.S.A., Brunham R.C., Nelson W.C., Paulsen I.T.,
 RA Heidelberg J.F., Holtzaple E.K., Khouri H.M., Federova N.B.,
 RA Carthy H.A., Unayam L.A., Haft D.H., Peterson J.D., Beanan M.J.,
 RA White O., Salzberg S.L., Hsia R.-C., McClarty G., Rank R.G.,
 RA Bavoll P.M., Fraser C.M.;
 RT "Genome sequence of Chlamydomophila caviae (Chlamydia psittaci GPIC):
 RT examining the role of niche-specific genes in the evolution of the
 RT Chlamydiaceae.";
 RL Nucleic Acids Res. 31:2134-2147(2003).
 RL EMBL; AR016996; AAP05280.1; -;
 DR TIGR; CCA00537; -;
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 140 AA; 15466 MW; 45345ECBDA1EB644 CRC64;

Query Match 100.0%; Score 32; DB 2; Length 140;
 Best Local Similarity 100.0%; Pred. No. 55;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
 |||||
 Db 66 CRGDD 70

RESULT 2
 Q8ZHJ9 ID Q8ZHJ9 PRELIMINARY; PRT; 174 AA.
 AC Q6ZHU9;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Putative Vacuolar ATP synthase subunit G 1.
 GN Name=OU1218_D07.1-1; Synonyms=OU111_E07.26-1;

OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OC NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP004052; RAD07544.1; -;
 DR EMBL; AP003994; RAD07530.1; -;
 DR InterPro; IPR009019; KH_prok.
 DR InterPro; IPR005124; V-ATPase_G.
 DR Pfam; PF03179; V-ATPase_G; 1.
 SQ SEQUENCE 174 AA; 19423 MW; D8EAD87DC7369580 CRC64;

Query Match 100.0%; Score 32; DB 2; Length 174;
 Best Local Similarity 100.0%; Pred. No. 69;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
 |||||
 Db 21 CRGDD 25

RESULT 3
 Q9LEC1 ID Q9LEC1 PRELIMINARY; PRT; 187 AA.
 AC Q9LEC1;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Inhibitor of yeast proteinase A (Fragment).
 OS Lycopersicon esculentum (Tomato)
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC Lamiids; Solanales; Solanaceae; Solanum.
 OC NCBI_TaxID=4081;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Leaf;
 RA Cater S.A.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ289776; CAB94854.1; -;
 DR GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
 DR InterPro; IPR011065; Kunitz like.
 DR InterPro; IPR002160; Prot_inh_Kunz-1g.
 DR Pfam; PF0197; Kunitz_legume; 1.
 DR PRINTS; PR00291; KUNITZINHPT.
 DR ProDom; PD000891; Prot_inh_Kunz-1g; 1.
 DR SMART; SM00452; STI; 1.
 DR PROSITE; PS00283; SOYBEAN_KUNITZ; 1.
 FT NON_TER
 SQ SEQUENCE 187 AA; 20525 MW; 3229395B366D1B2A CRC64;

Query Match 100.0%; Score 32; DB 2; Length 187;
 Best Local Similarity 100.0%; Pred. No. 74;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
 |||||
 Db 152 CRGDD 156

RESULT 4
 Q45047 ID Q45047 PRELIMINARY; PRT; 199 AA.
 AC Q45047;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Putative trypsin-like protein (Fragment).
 GN Name=TL3;

OS Scirpophaga incertulas.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Pyraloidea;
 OC Pyralidae; Schoenobiinae; Scirpophaga.
 OX NCBI_TaxID=72366;

RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=IRRI; TISSUE=Midgut;
 RC MEDLINE=20110891; PubMed=10646971; DOI=10.1016/S0965-1748(99)00097-1;
 RA Mazumdar-Leighton S., Babu C.R., Bennett J.;
 RT "Identification of novel serine proteinase gene transcripts in the
 RT midguts of two tropical insect pests, Scirpophaga incertulas (Wk.) and
 RT Helicoverpa armigera";
 RL Insect Biochem. Mol. Biol. 30:57-68(2000).
 RN [2]

RN SEQUENCE FROM N.A.
 RC STRAIN=IRRI; TISSUE=Midgut;
 RA Mazumdar S.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF045142; AAC02220.1; -;
 DR HSSP; P00760; 3BTH.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004295; F:trypsin activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR009003; Pept_Ser_Cys.
 DR Pfam; PF00089; Trypsin; 1.
 DR SMART; SM00020; Tryp_SPC; 1.
 DR PROSITE; PS0240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
 KW Hydrolase; Protease; Serine protease.
 FT NON_TER 1
 FT SIGNAL 1
 SQ SEQUENCE 199 AA; 21089 MW; 72E0DC50C0BB1968 CRC64;

Query Match 100.0%; Score 32; DB 2; Length 199;
 Best Local Similarity 100.0%; Pred. No. 79;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
 |||||
 Db 151 CRGDD 155

RESULT 5
 Q9MCC2 ID Q9MCC2 PRELIMINARY; PRT; 215 AA.
 AC Q9MCC2;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Vpf215.
 GN Names=vpf215;
 OS Bacteriophage vFO4K68.
 OC Viruses; ssDNA viruses; Inoviridae; Inovirus.
 OX NCBI_TaxID=127508;
 RN [1]
 RN SEQUENCE FROM N.A.
 RA Chang B., Miyamoto H., Taniguchi H., Yoshida S.;
 RT "Isolation and Genetic Characterization of a Novel Filamentous
 RT Bacteriophage, a Deleted Form of Phage f237, from a Pandemic Vibrio
 RT parahaemolyticus O4: K68 Strain."
 RL Microbiol. Immunol. 46:565-569(2002).
 DR EMBL; AB043679; BAA96438.1; -;
 SQ SEQUENCE 215 AA; 24626 MW; FF0839A0657EAD1B CRC64;

Query Match 100.0%; Score 32; DB 2; Length 215;
 Best Local Similarity 100.0%; Pred. No. 85;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
 |||||
 Db 97 CRGDD 101

RESULT 6
 Q9LEG1 ID Q9LEG1 PRELIMINARY; PRT; 220 AA.
 AC Q9LEG1;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Cathespin D inhibitor precursor.
 GN Name=cathDinh;
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC Lamiales; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4081;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Leaf;
 RA Lison P., Rodrigo I., Conejero V.;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ295638; CAC00536.1; -;
 DR GO; GO:0004866; F:endopeptidase inhibitor activity; IEA.
 DR InterPro; IPR011065; Kunitz like.
 DR InterPro; IPR002160; Prot_inh_Kunz_lg.
 DR Pfam; PF00197; Kunitz_legume; 1.
 DR PRINTS; PR00291; KUNITZINHBR.
 DR ProDom; PD000891; Prot_inh_Kunz_lg; 1.
 DR SMART; SM00452; STI; 1.
 DR PROSITE; PS00283; SOYBEAN_KUNITZ; 1.
 KW Signal.
 FT SIGNAL 1
 FT SIGNAL 20 Potential.
 SQ SEQUENCE 220 AA; 24188 MW; 5DFFC488938D1F6A CRC64;

Query Match 100.0%; Score 32; DB 2; Length 220;
 Best Local Similarity 100.0%; Pred. No. 87;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
 |||||
 Db 185 CRGDD 189

RESULT 7
 Q6D9R3 ID Q6D9R3 PRELIMINARY; PRT; 230 AA.
 AC Q6D9R3;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Hypothetical protein.
 GN Ordered locus names=ECA0552;
 OS Erwinia carotovora (subsp. atroseptica) (Pectobacterium atrosepticum).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Pectobacterium.
 OX NCBI_TaxID=29471;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=SCRI 1043 / ATCC BAA-672;
 RC PubMed=15263089; DOI=10.1073/pnas.0402424101;
 RX Bell K.S., Sebahia M., Pritchard L., Holden M.T.G., Hyman L.J.,
 RA Holeva M.C., Thomson N.R., Brooks K., Churcher L.J.C., Mungall K.,
 RA Atkin R., Bason N., Hauser H., Chillingworth T., Clark K., Doggett J.,
 RA Fraser A., Hance Z., Moule S., Mouton R., Norbertczak H.,
 RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
 RA Salmond G.P.C., Birch P.R.J., Parkhill J., Toth I.K.;
 RT "Genome sequence of the enterobacterial phytopathogen Erwinia
 RT carotovora subsp. atroseptica and characterization of virulence
 RT factors."
 RL Proc. Natl. Acad. Sci. U.S.A. 101:11105-11110(2004).
 DR EMBL; BX950851; CAG73467.1; -;
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 230 AA; 26285 MW; 7E8007B2AADB73F CRC64;

Query Match 100.0%; Score 32; DB 2; Length 230;
 Best Local Similarity 100.0%; Pred. No. 91;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
 DB 197 CRGDD 201

RESULT 8
 Y166_METJA
 ID Y166_METJA STANDARD; PRT; 255 AA.
 AC Q57630;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Hypothetical UPF0204 protein MJ0166.
 GN OrderedLocustNames=MJ0166;
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 OC Methanocaldococcaceae; Methanocaldococcus.
 OX NCBI_TaxID=2190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=8688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kervlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
 RA Usterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.:
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 jannaschii.";
 RL Science 273:1058-1073 (1996).
 CC -!- SIMILARITY: Belongs to the UPF0204 family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U67473; AAB98148.1; -.
 DR F1R; G64320; G64320.
 DR TIGR; MJ0166; -.
 DR HAMAP; MF 00562; -; 1.
 DR InterPro; IPR011007; B12 binding.
 DR InterPro; IPR007508; DUF516.
 DR Pfam; PF04414; DUF516; 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 255 AA; 29072 MW; 6428DF7AEC802CE4 CRC64;

Query Match 100.0%; Score 32; DB 1; Length 255;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
 DB 233 CRGDD 237

RESULT 9
 Q9ZWQ4
 ID Q9ZWQ4 PRELIMINARY; PRT; 280 AA.
 AC Q9ZWQ4;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE UDP-glycose:flavonoid glycosyltransferase (Fragment).
 GN Name=UG1YT;
 OS Vigna mungo (Rice bean) (Black gram).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Vigna.
 OX NCBI_TaxID=3915;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NIAS no. R-29; TISSUE=First simple leaf;
 RX MEDLINE=99108563; PubMed=9891414;
 RA Mato M., Ozeki Y., Itoh Y., Higeta D., Yoshitama K., Teramoto S.,
 RA Aida R., Ishikura N., Shibata M.;
 RT "Isolation and characterization of a cDNA clone of UDP-galactose:
 RT flavonoid 3-O-galactosyltransferase (UF3GaT) expressed in Vigna mungo
 RT seedlings.";
 RL Plant Cell Physiol. 39:1145-1155 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NIAS no. R-29; TISSUE=First simple leaf;
 RA Mato M., Sibata M., Ozeki Y.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: Belongs to the UDP-glycosyltransferase family.
 DR EMBL; AB012115; BAA36411.1; -.
 DR GO; GO:0016758; F:transferase activity, transferring hexosyl . . .; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR002213; UDP_glucos_trans.
 DR Pfam; PF00201; UDPGT; 1.
 DR PROSITE; PS00375; UDPGT; 1.
 KW Glycosyltransferase; Transferase.
 FT NON_TER 1
 SQ SEQUENCE 280 AA; 30582 MW; 0A4A3705550B55A0 CRC64;

Query Match 100.0%; Score 32; DB 2; Length 280;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
 DB 66 CRGDD 70

RESULT 10
 Q9X6W2
 ID Q9X6W2 PRELIMINARY; PRT; 289 AA.
 AC Q9X6W2;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Oligopeptide transport system permease protein (Fragment).
 GN Name=oppC;
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OG Plasmid megaplasmid 2.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CXMI-188;
 RA Krol E.A., Soberon M., Yurgel S.N., Miranda Rios J., Simarov B.V.;
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Part of a binding-protein-dependent transport system.
 CC Probably responsible for the translocation of the substrate across
 CC the membrane (By similarity).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -!- SIMILARITY: Belongs to the binding-protein-dependent transport
 CC system permease family.
 DR EMBL; AF148072; AAD33993.1; -.
 DR F1R; T44599; T44599.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005215; P:transporter activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR000515; BPD_transp.

```

DR InterPro: IPR011028; Cyclin-like.
DR Pfam: PF00528; BPD_transp_1; 1.
DR PROSITE: PS00928; ABC_TM1; 1.
KW Plasmid; Transmembrane; Transport.
FT NON TER 1
SQ SEQUENCE 289 AA; 31798 MW; F49623B724700D58 CRC64;

Query Match 100.0%; Score 32; DB 2; Length 289;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
Db 36 CRGDD 40

RESULT 11
HEM3_XANAC
ID HEM3_XANAC STANDARD; PRT; 304 AA.
AC Q8PPE3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Porphobilinogen deaminase (EC 2.5.1.61) (PBG) (Hydroxymethylbilane
DE synthase) (HMB) (Pre-uroporphyrinogen synthase).
GN Name-hemC; OrderedLocusNames=XAC0622;
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINES=2022145; PubMed=12024217; DOI=10.1038/417459a;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chamberg F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyana A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RA "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
CC -!- FUNCTION: Tetrapolymerization of the monopyrrole PBG into the
CC hydroxymethylbilane preuroporphyrinogen in several discrete steps.
CC -!- CATALYTIC ACTIVITY: 4 porphobilinogen + H(2)O =
CC hydroxymethylbilane + 4 NH(3).
CC -!- COFACTOR: Covalently binds a dipyrromethane cofactor to which the
CC porphobilinogen subunits are added (By similarity).
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- PATHWAY: Porphyrin biosynthesis by the C5 pathway; fourth step.
CC -!- SIMILARITY: Belongs to the HMB family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: A5011690; AAM35511.1; -.
CC HSP: P06983; 1YPN.
CC HAMAP: MF_00260; -. 1.

```

```

CC -----
DR EMBL; AB012470; AAM42781.1; -.
DR HSP; P06983; 1YPN.
DR HAMAP; MF 00260; -: 1.
DR InterPro; IPR000860; Porphobil_deam.
DR Pfam; PF01379; Porphobil_deam; 1.
DR Pfam; PF03900; Porphobil_deamC; 1.
DR PRINTS; PR00151; PORPHBDNASE.
DR ProDom; PD002745; Porphobil_deam; 1.
DR TIGRFAMs; TIGR00212; hemC; 1.
DR PROSITE; PS00533; PORPHOBILINOGEN_DEAM; 1.
KW Complete proteome; Porphyrin biosynthesis; Transferase.
FT BINDING 240 240 Pyrimethane cofactor (By similarity).
SQ SEQUENCE 304 AA; 32460 MW; F77A49FB1932FF02 CRC64;

Query Match 100.0%; Score 32; DB 1; Length 304;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
Db 203 CRGDD 207

RESULT 13
LPP3 HUMAN STANDARD; PRT; 311 AA.
AC O14495; O96GW0; Q99782;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DE 05-JUL-2004 (Rel. 44, Last annotation update)
DE Lipid phosphate phosphohydrolase 3 (EC 3.1.3.4) (Phosphatidic acid
DE phosphate 2b) (Phosphatidate phosphohydrolase type 2b) (PAP2b) (PAP-
DE 2b) (PAP2-beta) (Vascular endothelial growth factor and type I
DE collagen inducible protein) (VCIP).
GN Name=PPAP2B; Synonyms=LPP3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A., CHARACTERIZATION, AND INDUCTION.
RX MEDLINE=97450990; PubMed=9305923; DOI=10.1074/jbc.273.39.24572;
RA Kai M., Wada I., Imai S.-I., Sakane F., Kanoh H.;
RT "Cloning and characterization of two human isozymes of Mg2+-
RT independent phosphatidic acid phosphatase.";
RL J. Biol. Chem. 273:24572-24578 (1997).
[2]
RN SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=98371049; PubMed=9705349; DOI=10.1074/jbc.273.34.22059;
RA Roberts R., Sciorra V.A., Morris A.J.;
RT "Human type 2 phosphatidic acid phosphohydrolases. Substrate
RT specificity of the type 2a, 2b, and 2c enzymes and cell surface
RT activity of the 2a isoform.";
RL J. Biol. Chem. 273:22059-22067 (1998).
[3]
RN SEQUENCE FROM N.A.
RX PubMed=12660161; DOI=10.1093/emboj/cdg165;
RA Humceoe J.O., Feng S., Thakker G.D., Yang J., Hong J., Wary K.K.;
RT "Regulation of cell-cell interactions by phosphatidic acid phosphatase
RT 2b/VCIP.";
RL EMBO J. 22:1539-1554 (2003).
[4]
RN SEQUENCE FROM N.A.
RX Leung D.W., Tompkins C.K.;
RT "Molecular cloning of and expression of an isoform of human
RT phosphatidic acid phosphatase cDNA.";
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
[5]
RN SEQUENCE FROM N.A.
RX MEDLINE=97264341; PubMed=9110174;
RA Yu W., Anderson B., Worley K.C., Muzny D.M., Ding Y., Liu W.,
RA Ricafrente J.Y., Wentland M.C., Lennon G., Gibbs R.A.;

```

```

RT "Large-scale concatenation cDNA sequencing.";
RL Genome Res. 7:353-358 (1997).
RN [6]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Klausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[7]
RN SUBUNIT.
RX PubMed=14725715; DOI=10.1186/1471-2091-5-2;
RA Burnett C., Makridou P., Hewlett L., Howard K.;
RT "Lipid phosphate phosphatases dimerise, but this interaction is not
RT required for in vivo activity.";
RL BMC Biochem. 5:2-2 (2004).
CC -!- FUNCTION: Catalyzes the conversion of phosphatidic acid (PA) to
CC diacylglycerol (DG). In addition it hydrolyzes lysophosphatidic
CC acid (LPA), ceramide-1-phosphate (C-1-P) and sphingosine-1-
CC phosphate (S-1-P). The relative catalytic efficiency is LPA = PA >
CC C-1-P > S-1-P. May be involved in cell adhesion and in cell-cell
CC interactions.
CC -!- CATALYTIC ACTIVITY: A 3-sn-phosphatidate + H(2)O = 1,2-diacyl-
CC sn-glycerol + phosphate.
CC -!- ENZYME REGULATION: Inhibited by sphingosine, zinc ions and
CC propanolol. Not inhibited by N-ethylmaleimide treatment.
CC -!- SUBUNIT: Homodimer. This complex seems not to be involved in
CC substrate recognition, it may confer only structural or functional
CC stability.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Post-Golgi and
CC plasma membrane localization.
CC -!- TISSUE SPECIFICITY: Ubiquitously expressed. Highly expressed in
CC heart and placenta.
CC -!- INDUCTION: By epidermal growth factor (EGF), vascular endothelial
CC growth factor (VEGF), basic fibroblast growth factor (bFGF) and
CC phorbol myristate acetate (PMA).
CC -!- PTM: N-glycosylated. Contains high-mannose oligosaccharides.
CC -!- SIMILARITY: Belongs to the PA-phosphatase related phosphoesterase
CC family.
CC -!- CAUTION: Ref 5 sequence differs from that shown due to a
CC frameshift in position 225.
CC -----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB000889; BAA22594.1; -.
DR EMBL; AF017786; AAC63433.1; -.
DR EMBL; AF480883; AAO84481.1; -.
DR EMBL; AF043329; AAO22271.1; -.
DR EMBL; U79294; AAB50222.1; ALT_FRAME.
DR EMBL; BC009196; AAB09196.1; -.

```


"generation and initial analysis of more than 15,000 full-length human RT
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RN SEQUENCE FROM N.A.
RP
RC TISSUE=Testis;
RA Director MGC Project;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC070117; AAH70117.1; -.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; Ank; 4.
DR PRINTS; Prol415; ANKYRN.
DR SMART; SM00248; ANK; 4.
DR PROSITE; PS50088; ANK_REPEAT; 3.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
KW ANK repeat.
SQ SEQUENCE 321 AA; 35410 MW; 60C46EFC62716EB4 CRC64;

Query Match 100.0%; Score 32; DB 2; Length 321;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
|||
Db 25 CRGDD 29

RESULT 15
Q820U7 PRELIMINARY; PRT; 357 AA.
ID Q820U7 Q7C4T9;
AC Q820U7; Q7C4T9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein STY4914.
GN OrderedLocusNames=STY4914, t4606;
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA Baker S., Basham D., Brooks K., Chillingworth I., Conerton P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Felwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RT "complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
[2]
RN SEQUENCE FROM N.A.
RP
RC STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RL DOI=10.1128/JB.185.7.2330-2337.2003;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.P., Rose D.J.,
RA Burland V., Kodoyianni V., Schwartz D.C., Blattner P.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18";
RL J. Bacteriol. 185:2330-2337(2003).
DR EMBL; AL627284; CAD03398.1; -.
DR EMBL; AB016849; AAO72038.1; -.
DR GO; GO:0003824; P:catalytic activity; IEA.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR Pfam; PF01734; Patatin; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 357 AA; 39633 MW; A3432EC13D097E6B CRC64;

"generation and initial analysis of more than 15,000 full-length human RT
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RN SEQUENCE FROM N.A.
RP
RC TISSUE=Testis;
RA Director MGC Project;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC070117; AAH70117.1; -.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; Ank; 4.
DR PRINTS; Prol415; ANKYRIN.
DR SMART; SM00248; ANK; 4.
DR PROSITE; PS50088; ANK_REPEAT; 3.
DR PROSITE; PS50297; ANK_REP_REGION; 1.
KW ANK repeat.
SQ SEQUENCE 321 AA; 35410 MW; 60C46EFC62716EB4 CRC64;

Query Match 100.0%; Score 32; DB 2; Length 321;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CRGDD 5
|||
Db 25 CRGDD 29

RESULT 15

Q820U7 PRELIMINARY; PRT; 357 AA.

ID Q820U7 Q7C4T9;
AC Q820U7; Q7C4T9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein STY4914.
GN OrderedLocusNames=STY4914, t4606;
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA Baker S., Basham D., Brooks K., Chillingworth I., Conerton P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Felwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RT "complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
[2]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RL DOI=10.1128/JB.185.7.2330-2337.2003;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.P., Rose D.J.,
RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18";
RL J. Bacteriol. 185:2330-2337(2003).
DR EMBL; AL627284; CAD03398.1; -.
DR EMBL; AB016849; AAO72038.1; -.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR Pfam; PF01734; Patatin; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 357 AA; 39633 MW; A3432EC13D097E6B CRC64;

```

Query Match      100.0%; Score 32; DB 2; Length 357;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CRGDD 5
DB      139 CRGDD 143

RESULT 16
Q8ZJW2
ID      Q8ZJW2      PRELIMINARY;      PRT;      357 AA.
AC      Q8ZJW2;
DT      01-NAR-2002 (TrEMBLrel. 20, Created)
DT      01-NAR-2002 (TrEMBLrel. 20, Last sequence update)
DT      01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE      Putative phosphoesterase.
GN      Name=yjjj; OrderedLocusNames=STW4563;
OS      Salmonella typhimurium.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC      Enterobacteriaceae; Salmonella.
OX      NCBI_TaxID=602;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=LT2;
RX      MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;
RA      McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA      Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA      Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA      Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA      Waterston R., Wilson R.K.;
RT      "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT      LT2.";
RL      Nature 413:852-856 (2001).
DR      EMBL; AB008914; AAL23378.1; -.
DR      GO; GO:0003824; F:catalytic activity; IEA.
DR      GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR      Pfam; PF01734; Patatin; 1.
KW      Complete proteome.
SQ      SEQUENCE 357 AA; 39650 MW; 945E297E5DB3394E CRC64;

Query Match      100.0%; Score 32; DB 2; Length 357;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CRGDD 5
DB      139 CRGDD 143

RESULT 17
Q92VE0
ID      Q92VE0      PRELIMINARY;      PRT;      363 AA.
AC      Q92VE0;
DT      01-DEC-2001 (TrEMBLrel. 19, Created)
DT      01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT      01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Putative mureinpeptideoligopeptide ABC transporter permease
DE      protein.
GN      ORFNames=SWb21263;
GN      Rhizobium meliloti (Sinorhizobium meliloti).
OG      Plasmid pSymB.
OC      Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC      Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX      NCBI_TaxID=382;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=1021;
RX      MEDLINE=21396508; PubMed=11481431; DOI=10.1073/pnas.161294698;
RA      Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
RA      Vorhoeelter F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
RA      Golding B., Puehler A.;
RT      "The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-
```

```

RT      fixing endosymbiont Sinorhizobium meliloti.";
RT      Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894 (2001).
CC      -!- FUNCTION: Part of a binding-protein-dependent transport system.
CC      Probably responsible for the translocation of the substrate across
CC      the membrane (By similarity).
CC      -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC      -!- SIMILARITY: Belongs to the binding-protein-dependent transport
CC      system permease family.
DR      EMBL; AL591985; CAC49167.1; -.
DR      PIR; G95937; G95937.
DR      GO; GO:0016021; C:integral to membrane; IEA.
DR      GO; GO:0005215; F:transporter activity; IEA.
DR      GO; GO:0006810; F:transport; IEA.
DR      InterPro; IPR000515; BPD transp.
DR      InterPro; IPR011028; Cyclin like.
DR      Pfam; PF00528; BPD transp. 1.
DR      PROSITE; PS50928; ABC_TWI_1.
KW      Complete proteome; Plasmid; Transmembrane; Transport.
SQ      SEQUENCE 363 AA; 40408 MW; 45718411C8821F24 CRC64;

Query Match      100.0%; Score 32; DB 2; Length 363;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CRGDD 5
DB      110 CRGDD 114

RESULT 18
Q98IE7
ID      Q98IE7      PRELIMINARY;      PRT;      372 AA.
AC      Q98IE7;
DT      01-OCT-2001 (TrEMBLrel. 18, Created)
DT      01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT      01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE      NADH-dependent dyhydrogenase.
GN      OrderedLocusNames=mlr2436;
OS      Rhizobium loti (Mesorhizobium loti).
OC      Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC      Phyllobacteriaceae; Mesorhizobium.
OX      NCBI_TaxID=381;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=MAFF303099;
RX      MEDLINE=21082930; PubMed=11214968;
RA      Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA      Watanabe A., Iidesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA      Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA      Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
RA      Takeuchi C., Yamada M., Tabata S.;
RT      "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT      Mesorhizobium loti.";
RL      DNA Res. 7:331-338 (2000).
DR      EMBL; AP002999; BAB49569.1; -.
DR      GO; GO:0016491; F:oxidoreductase activity; IEA.
DR      GO; GO:0006118; P:electron transport; IEA.
DR      InterPro; IPR000683; GFO/IDH/MocA_N.
DR      Pfam; PF01408; GFO_IDH_MocA; 1.
KW      Complete proteome.
SQ      SEQUENCE 372 AA; 40635 MW; 1A1A307A237B8CCD CRC64;

Query Match      100.0%; Score 32; DB 2; Length 372;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CRGDD 5
DB      167 CRGDD 171

RESULT 19
Q8S9A8
```

```

ID Q8S9A8 PRELIMINARY; PRT; 390 AA.
AC Q8S9A8;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Glucosyltransferase-1 (Fragment).
GN Name=AGT-1;
OS Phaseolus angularis (Adzuki bean) (Vigna angularis).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Vigna.
OX NCBI_TaxID=3914;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hypocotyl;
RU Xu Z.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the UDP-glycosyltransferase family.
DR EMBL; AB070743; BAB86919.1; -.
DR GO; GO:0016758; F:transferase activity, transferring hexosyl . . . IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002213; UDP-glucos_trans.
DR Pfam; PF00201; UDPGT; 1.
DR PROSITE; PS00375; UDPGT; 1.
KW Glycosyltransferase; Transferase.
FT NON_TER 1
SQ SEQUENCE 390 AA; 43122 MW; AABDE97E71A8D009 CRC64;

Query Match 100.0%; Score 32; DB 2; Length 390;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
Db 176 CRGDD 180

RESULT 20
Q8LMH4 PRELIMINARY; PRT; 393 AA.
ID Q8LMH4;
AC Q8LMH4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein OSUNBA0032N04.16;
GN Name=OSUNBA0032N04.16;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA McCombie W.R., de la Bastide M., Spiegel L., Preston R., Ferraro K.,
RA Kuit K., Nascimento L., Zutavern T., Ballia V., Bell M., Baker J.,
RA Santos L., Miller B., Katzenberger F., Muller S., King L., Yang C.,
RA Dike S., O'Shaughnessy A., Palmer L., Dedhia N.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC114474; AAM74325.2; -.
DR Gramene; Q8LMH4; -.
DR InterPro; IPR001810; F-box.
DR Pfam; PF00846; F-box; 1.
DR SMART; SM00256; FBOX; 1.
KW Hypothetical protein.
SQ SEQUENCE 393 AA; 45232 MW; 7C85AB25EF0A70C2 CRC64;

Query Match 100.0%; Score 32; DB 2; Length 393;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
Db 175 CRGDD 179

```

```

RESULT 21
PNCB ECO57 STANDARD; PRT; 399 AA.
ID PNCB ECO57 STANDARD; PRT; 399 AA.
AC Q8XD88;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Nicotinate phosphoribosyltransferase (EC 2.4.2.11) (NAPRTase).
GN Name=pncB; OrderedLocusNames=21279, EGS1014;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.P., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoudis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RL Nature 409:529-533 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / Sakai / RIMD 0509952 / EHEC;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:11-22 (2001).
CC -!- CATALYTIC ACTIVITY: Nicotinate D-ribonucleotide + diphosphate =
CC nicotinate + 5-phospho-alpha-D-ribose 1-diphosphate.
CC -!- PATHWAY: NAD biosynthesis; nicotinamide to NAMN; second step.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the NAPRTase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB005283; AAG55416.1; -.
CC EMBL; AP002553; BAB34437.1; -.
CC PIR; D85619; D85619.
CC PIR; F90755; F90755.
CC HAMAP; MF_00570; -.
CC InterPro; IPR007229; NAPRTase.
CC InterPro; IPR006406; Nic Prtrans.
CC InterPro; IPR008967; P53_like_DNA_bnd.
CC Pfam; PF04095; NAPRTase; 1.
CC PIRSF; PIRSF000484; Nicot_phos_ribo; 1.
CC TIGRFAMs; TIGR01514; NAPRTase; 1.
KW Complete proteome; Glycosyltransferase;
KW Pyridine nucleotide biosynthesis; Transferase.
FT INIT_MET 0 By similarity
SQ SEQUENCE 399 AA; 45838 MW; 47DAC387A0EC6926 CRC64;

Query Match 100.0%; Score 32; DB 1; Length 399;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5

```

```

Db          40 CRGDD 44

RESULT 22
PNCB_ECOLI
ID PNCB_ECOLI STANDARD; PRT; 399 AA.
AC P18133;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Nicotinate phosphoribosyltransferase (EC 2.4.2.11) (NAPRTase).
GN Name=pncB; OrderedLocusNames=b0931;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-20.
RC STRAIN=GEC70;
RX MEDLINE=91009224; PubMed=2211655;
RA Rubolts M.G., Terpstra P., van Beilen J.B., Kingma J.,
RA Meesters H.A.R., Witholt B.;
RT "Variation of cofactor levels in Escherichia coli. Sequence analysis
RT and expression of the pncB gene encoding nicotinic acid
RT phosphoribosyltransferase."
RL J. Biol. Chem. 265:17665-17672(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-Kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map."
RL DNA Res. 3:137-155(1996).
RN [4]
RP CATALYTIC ACTIVITY: Nicotinate D-ribonucleotide + diphosphate =
CC nicotinate + 5-phospho-alpha-D-ribose 1-diphosphate.
CC -!- PATHWAY: NAD biosynthesis; nicotinamide to NADH; second step.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- MISCELLANEOUS: Requires ATP as well as phosphoribosyl
CC pyrophosphate for activity.
CC -!- SIMILARITY: Belongs to the NAPRTase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J05568; AAA24400.1; -.
CC EMBL; U00096; AAC74017.1; -.
CC EMBL; D90731; BAA35683.1; -.
CC PIR; JQ0756; JQ0756.
CC ECHOBASE; EB0735.
CC EcoGene; EG10742; pncB.
CC HAMAP; MF_00570; -; 1.

```

```

DR InterPro; IPR007229; NAPRTase.
DR InterPro; IPR006406; Nic Ptrans.
DR InterPro; IPR008967; P53_like_DNA_bnd.
DR Pfam; PF04095; NAPRTase; 1.
DR PIRSF; PIRSF000484; Nicot_phos_ribo; 1.
DR TIGRFAMs; TIGR01514; NAPRTase; 1.
KW Complete proteome; Direct protein sequencing; Glycosyltransferase;
KW Pyridine nucleotide biosynthesis; Transferase.
FT INIT MET 0
SQ SEQUENCE 399 AA; 45766 MW; 37F0C0D2C3BA5C45 CRC64;

Query Match 100.0%; Score 32; DB 1; Length 399;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
DB 40 CRGDD 44

RESULT 23
PNCB_SALTI
ID PNCB_SALTI STANDARD; PRT; 399 AA.
AC Q827Y9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Nicotinate phosphoribosyltransferase (EC 2.4.2.11) (NAPRTase).
GN Name=pncB; OrderedLocusNames=STY1010, t1930;
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18."
RL Nature 413:848-852(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TY2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RX DOI=10.1128/JB.185.7.2330-2337.2003;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyanni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18."
RL J. Bacteriol. 185:2330-2337(2003).
RN [3]
RP CATALYTIC ACTIVITY: Nicotinate D-ribonucleotide + diphosphate =
CC nicotinate + 5-phospho-alpha-D-ribose 1-diphosphate.
CC -!- PATHWAY: NAD biosynthesis; nicotinamide to NADH; second step.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the NAPRTase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL627268; CAD05404.1; -.

```

```

DR EMBL; AE016840; AAC69546.1; -.
DR HAMAP; MF_00570; -, 1.
DR InterPro; IPR007229; NAPTase.
DR InterPro; IPR006406; Nic_Ptrans.
DR Pfam; PF04095; NAPTase; 1.
DR PIRSF; PIRSF000484; Nicot_phos_ribo; 1.
DR TIGRFAMs; TIGR01514; NAPTase; 1.
DR PIRSF; PIRSF000484; Nicot_phos_ribo; 1.
DR TIGRFAMs; TIGR01514; NAPTase; 1.
DR Complete proteome; Glycosyltransferase;
KW Pyridine nucleotide biosynthesis; Transferase.
FT INIT MET 0
SQ SEQUENCE 399 AA; 45562 MW; 45626 MW; ADF38AEE4618214 CRC64;

Query Match 100.0%; Score 32; DB 1; Length 399;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
DB 40 CRGDD 44

RESULT 24
PNCB SALTY STANDARD; PRT; 399 AA.
ID PNCB SALTY STANDARD; PRT; 399 AA.
AC P22253; 1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Nicotinate phosphoribosyltransferase (EC 2.4.2.11) (NAPTase).
GN Name=pncB; OrderedLocusNames=STW1004;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91100340; PubMed=1987148;
RA Vinitsky A., Teng H., Grubmeyer C.T.;
RT "Cloning and nucleic acid sequence of the Salmonella typhimurium pncB
RT gene and structure of nicotinate phosphoribosyltransferase.";
RL J. Bacteriol. 173:536-540(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=J2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=1167609; DOI=10.1038/35101614;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
CC -1- CATALYTIC ACTIVITY: Nicotinate D-ribose nucleotide + diphosphate =
CC nicotinate + 5-phospho-alpha-D-ribose 1-diphosphate.
CC -1- PATHWAY: NAD biosynthesis; nicotinamide to NAMN; second step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the NAPTase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M55986; AAA27190.1; -.
DR EMBL; AE008743; AAL19938.1; -.
DR PIR; A39130; A39130.
DR StGene; SG10305; pncB.
DR HAMAP; MF_00570; -, 1.

```

```

DR InterPro; IPR007229; NAPTase.
DR InterPro; IPR006406; Nic_Ptrans.
DR Pfam; PF04095; NAPTase; 1.
DR PIRSF; PIRSF000484; Nicot_phos_ribo; 1.
DR TIGRFAMs; TIGR01514; NAPTase; 1.
DR Complete proteome; Glycosyltransferase;
KW Pyridine nucleotide biosynthesis; Transferase.
FT INIT MET 0
SQ SEQUENCE 399 AA; 45530 MW; 90D9C3EDD8C092A6 CRC64;

Query Match 100.0%; Score 32; DB 1; Length 399;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
DB 40 CRGDD 44

RESULT 25
Q7UD27
ID Q7UD27 PRELIMINARY; PRT; 400 AA.
AC Q7UD27; 2003 (TREMELrel. 25, Created)
DT 01-OCT-2003 (TREMELrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 26, Last annotation update)
DE Nicotinate phosphoribosyltransferase.
GN Name=pncB; OrderedLocusNames=S0992;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2457T;
RX MEDLINE=22590274; PubMed=12704152;
RX DOI=10.1128/JAI.71.5.2775-2786.2003;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. II, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner P.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T.";
RL Infect. Immun. 71:2775-2786(2003).
DR EMBL; AE016981; AAP16443.1; -.
DR GO; GO:0004516; F:nicotinate phosphoribosyltransferase activity; IEA.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl...; IEA.
DR GO; GO:0019357; P:nicotinate nucleotide biosynthesis; IEA.
DR GO; GO:0019363; P:pyridine nucleotide biosynthesis; IEA.
DR InterPro; IPR007229; NAPTase.
DR InterPro; IPR006406; Nic_Ptrans.
DR Pfam; PF04095; NAPTase; 1.
DR PIRSF; PIRSF000484; Nicot_phos_ribo; 1.
DR TIGRFAMs; TIGR01514; NAPTase; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 400 AA; 45911 MW; 5385BCB1A8703095 CRC64;

Query Match 100.0%; Score 32; DB 2; Length 400;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
DB 41 CRGDD 45

RESULT 26
Q83LN3
ID Q83LN3 PRELIMINARY; PRT; 416 AA.
AC Q83LN3.
DT 01-JUN-2003 (TREMELrel. 24, Created)

```

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Nicotinate phosphoribosyltransferase.
 GN Name=pncB; OrderedLocusNames=SF0928;
 OS Shigella flexneri.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Shigella.
 OX NCBI_TaxID=623;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=301 / Serotype 2a;
 RX MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
 RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
 RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
 RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
 RA Yu J.;
 RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
 through comparison with genomes of Escherichia coli K12 and O157.";
 RL Nucleic Acids Res. 30:4432-4441(2002).
 DR EMBL; AE015122; AAN42557.1; -;
 DR GO; GO:004516; F:nicotinate phosphoribosyltransferase activity; IEA.
 DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
 DR GO; GO:0019357; P:nicotinate nucleotide biosynthesis; IEA.
 DR GO; GO:0019363; P:pyridine nucleotide biosynthesis; IEA.
 DR InterPro; IPR007229; NAPRTase.
 DR InterPro; IPR006406; Nic_Prtrans.
 DR InterPro; IPR008967; P53_like_DNA_bnd.
 DR Pfam; PF04095; NAPRTase; 1.
 DR TIGRFAMs; TIGR01514; NAPRTase; 1.
 DR Complete proteome; Glycosyltransferase; Transferase.
 KW Complete proteome; Glycosyltransferase; Transferase.
 SQ SEQUENCE 416 AA; 47700 MW; E43D723542AF2DD5 CRC64;
 Query Match 100.0%; Score 32; DB 2; Length 416;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CRGDD 5
 DB 57 CRGDD 61

RESULT 27
 Q8FJ98 PRELIMINARY; PRT; 416 AA.
 AC Q8FJ98;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Nicotinate phosphoribosyltransferase (EC 2.4.2.11).
 GN Name=pncB; OrderedLocusNames=c1073;
 OS Escherichia coli O6.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=217992;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=06.H1 / CFT073 / ATCC 700928 / UPEC;
 RX MEDLINE=22388334; PubMed=12471157; DOI=10.1073/pnas.252529799;
 RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
 RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
 RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
 RT "Extensive mosaic structure revealed by the complete genome sequence
 of uropathogenic Escherichia coli.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
 DR EMBL; AE016758; AAN79541.1; -;
 DR GO; GO:004516; F:nicotinate phosphoribosyltransferase activity; IEA.
 DR GO; GO:0019357; P:nicotinate nucleotide biosynthesis; IEA.
 DR GO; GO:0019363; P:pyridine nucleotide biosynthesis; IEA.
 DR InterPro; IPR007229; NAPRTase.
 DR InterPro; IPR006406; Nic_Prtrans.

DR InterPro; IPR008967; P53_like_DNA_bnd.
 DR Pfam; PF04095; NAPRTase; 1.
 DR TIGRFAMs; TIGR01514; NAPRTase; 1.
 DR Complete proteome.
 KW Complete proteome.
 SQ SEQUENCE 416 AA; 47722 MW; F00AFC07B04B1E23 CRC64;
 Query Match 100.0%; Score 32; DB 2; Length 416;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CRGDD 5
 DB 57 CRGDD 61

RESULT 28
 Q7XF76 PRELIMINARY; PRT; 428 AA.
 AC Q7XF76;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=OSJNB0032N04.17;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA The Rice Chromosome 10 Sequencing Consortium;
 RT "In-depth view of structure, activity, and evolution of rice
 chromosome 10.";
 RL Science 300:1566-1569(2003).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE017087; AAP53454.1; -;
 DR Gramene; Q7XF76; -;
 DR InterPro; IPR001810; F-box.
 DR Pfam; PF00646; F-box; 1.
 DR SMART; SM00256; FBOX; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 428 AA; 49415 MW; B2A9D80D969D2C1F CRC64;
 Query Match 100.0%; Score 32; DB 2; Length 428;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CRGDD 5
 DB 175 CRGDD 179

RESULT 29
 SYS_SYNY3 STANDARD; PRT; 430 AA.
 ID SYS_SYNY3
 AC P73201;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Seryl-tRNA synthetase (EC 6.1.1.11) (Serine--tRNA ligase) (SerRS).
 GN Name=serS; OrderedLocusNames=srl1703;
 OS Synechocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 OX NCBI_TaxID=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97061201; PubMed=8905231;
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,

RA Miyajima N., Hirose M., Sugita M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimizu S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
CC -1- CATALYTIC ACTIVITY: ATP + L-serine + tRNA(Ser) = AMP + diphosphate
CC + L-seryl-tRNA(Ser).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the class-II aminoacyl-tRNA synthetase
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D90904; BAA17227.1; -.
CC PIR: S75313; S75313.
CC HSP: P34945; ISEK.
CC HAMAP: MF_00176; -; 1.
CC InterPro: IPR002314; tRNA-synt_2b.
CC InterPro: IPR002317; tRNA-synt_ser.
CC InterPro: IPR010978; tRNA binding arm.
CC InterPro: IPR006195; tRNA_ligase_II.
CC Pfam: PF02403; Seryl_tRNA_N; 1.
CC Pfam: PF00587; tRNA-synt_2b; 1.
CC PRINTS: PR00981; TRNASYNTHSER.
CC TIGRFAMs: TIGR00414; serS; 1.
CC PROSITE: PS50862; AA tRNA LIGASE II; 1.
KW Aminoacyl-tRNA synthetase; ATP-binding; Complete proteome; Ligase;
KW Protein biosynthesis.
SQ SEQUENCE 430 AA; 48038 MW; 0E6A54434DB355E4 CRC64;

Query Match 100.0%; Score 32; DB 1; Length 430;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
Db 228 CRGDD 232

RESULT 30
Q804X1 PRELIMINARY; PRT; 442 AA.
AC Q804X1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Coagulation factor Vb (EC 3.4.21.21).
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RA Davidson C.J., Hirt R.P., Lal K., Snell P., Elgar G.,
RA Tuddenham E.G.D., McVey J.H.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to peptidase family S1.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
DR EMBL: AF465274; AAO33369.1; -.
DR HSP: P00740; 1CFH.
DR GO: GO:0005576; C:extracellular; IEA.

DR GO: GO:0005509; F:calcium ion binding; IEA.
DR GO: GO:0004263; F:chymotrypsin activity; IEA.
DR GO: GO:0003802; F:coagulation factor Vb activity; IEA.
DR GO: GO:0008233; F:peptidase activity; IEA.
DR GO: GO:0004295; F:trypsin activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR006209; EGF_like.
DR InterPro: IPR002383; GLA blood.
DR InterPro: IPR006210; IEGF.
DR InterPro: IPR001254; Peptidase S1.
DR InterPro: IPR001314; Peptidase_S1A.
DR InterPro: IPR009003; Pept_Ser_Cys.
DR InterPro: IPR000294; Vitk_dep_GLA.
DR Pfam: PF00008; EGF; 2.
DR Pfam: PF00594; Gla; 1.
DR PRINTS: PR00089; Trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00001; GLABLOOD.
DR SMART: SM00181; EGF; 2.
DR SMART: SM00069; GLA; 1.
DR SMART: SM00020; Tryp_SPC; 1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; 2.
DR PROSITE: PS50026; EGF_3; 1.
DR PROSITE: PS00011; GLA_1; 1.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
KW EGF-like domain; Hydrolase; Protease; Serine protease.
SQ SEQUENCE 442 AA; 48787 MW; 810A561A127F0CF7 CRC64;

Query Match 100.0%; Score 32; DB 2; Length 442;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
Db 383 CRGDD 387

RESULT 31
Q8S6K3 PRELIMINARY; PRT; 446 AA.
AC Q8S6K3;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein OSJNBa0019N10.23.
GN Name=OSJNBa0019N10.23;
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nipponbare;
RA McCombie W.R., de la Bastide M., Spiegel L., Preston R., Kirchoff K.,
RA Kuit K., Nascimento L., Zucavern T., Balija V., Bell M., Baker J.,
RA Santos L., Miller B., Katzenberger F., Muller S., King L., Yang C.,
RA O'Shaughnessy A., Palmer L., Dedhia N.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC092748; AAM01085.1; -.
DR Gramene; Q8S6K3; -.
DR InterPro: IPR001810; F-box.
DR Pfam: PF00646; F-box; 1.
DR SMART: SM00256; FBOX; 1.
KW Hypothetical protein.
SQ SEQUENCE 446 AA; 51647 MW; 361F9047C4CCA649 CRC64;

Query Match 100.0%; Score 32; DB 2; Length 446;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 CRGDD 5
Db      175 CRGDD 179

RESULT 32
Q8H4V4
ID      Q8H4V4      PRELIMINARY;      PRT;      446 AA.
AC      Q8H4V4;
DT      01-MAR-2003 (TrEMBLrel. 23, Created)
DT      01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Putative wall-associated kinase.
GN      Name=OJ1484.G09.127;
OS      Oryza sativa (japonica cultivar-group).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC      Ehrhartoideae; Oryzeae; Oryza.
OX      NCBI_TaxID=39947;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Sasaki T., Matsumoto T., Yamamoto K.;
RL      Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF003913; BAC24910.1; -.
DR      Gramene; Q8H4V4; -.
DR      GO; GO:0016301; F:kinase activity; IEA.
DR      InterPro; IPR006210; IEGF.
DR      SMART; SM00181; EGF; 1.
KW      Kinase.
SQ      SEQUENCE 446 AA; 47726 MW; 7BE3D91AD2ABC49C CRC64;

Query Match      100.0%; Score 32; DB 2; Length 446;
Best Local Similarity 100.0%; Pred.No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CRGDD 5
Db      232 CRGDD 236

RESULT 33
Q8VUB7
ID      Q8VUB7      PRELIMINARY;      PRT;      448 AA.
AC      Q8VUB7;
DT      01-MAR-2002 (TrEMBLrel. 20, Created)
DT      01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT      01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE      Hypothetical protein orf448.
GN      Name=orf448;
OS      Bradyrhizobium sp. ORS278.
OC      Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC      Bradyrhizobiaceae; Bradyrhizobium.
OX      NCBI_TaxID=114615;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      STRAIN=ORS278;
RX      MEDLINE=20570559; PubMed=11114184; DOI=10.1073/pnas.250484097;
RA      Giraud E., Hannibal L., Fardoux J., Jaubert M., Jourand P.,
RT      "Effect of Bradyrhizobium photosynthesis on stem nodulation of
RT      Aeschynomene sensitiva.";
RL      Proc. Natl. Acad. Sci. U.S.A. 97:14795-14800 (2000).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=ORS278;
RX      PubMed=14734565; DOI=10.1074/jbc.M312113200;
RA      Giraud E., Hannibal L., Fardoux J., Jaubert M., Jourand P.,
RA      Dreyfus B., Sturgis J.N., Vermeglio A.;
RT      "Two distinct crt gene clusters for two different functional classes
RT      of carotenoid in bradyrhizobium.";
RL      J. Biol. Chem. 279:15076-15083 (2004).
RN      [3]
RP      SEQUENCE FROM N.A.

```

```

RC      STRAIN=ORS278;
RA      Giraud E., Hannibal L., Fardoux J., Jaubert M., Vermeglio A.,
RA      Pignol D.;
RT      "Light and redox control of photosynthesis gene expression in
RT      Bradyrhizobium: Dual roles of two Ppsk.";
RL      Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF182374; AAL68699.1; -.
DR      GO; GO:0003677; F:DNA binding; IEA.
DR      GO; GO:0004803; F:transposase activity; IEA.
DR      GO; GO:0006313; P:DNA transposition; IEA.
DR      Pfam; PF01609; Transposase_11; 1.
SQ      SEQUENCE 448 AA; 50151 MW; 9387B4BB95BAA0FE CRC64;

Query Match      100.0%; Score 32; DB 2; Length 448;
Best Local Similarity 100.0%; Pred.No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CRGDD 5
Db      337 CRGDD 341

RESULT 34
Q6Z4U6
ID      Q6Z4U6      PRELIMINARY;      PRT;      469 AA.
AC      Q6Z4U6;
DT      05-JUL-2004 (TrEMBLrel. 27, Created)
DT      05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT      05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE      Putative Anthocyanin 5-aromatic acyltransferase.
GN      Name=OSJNBa0054L03.25;
OS      Oryza sativa (japonica cultivar-group).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC      Ehrhartoideae; Oryzeae; Oryza.
OX      NCBI_TaxID=39947;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Sasaki T., Matsumoto T., Katayose Y.;
RT      "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 8, BAC
RT      clone:OSJNBa0054L03.";
RL      Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AP005164; BAD05543.1; -.
DR      GO; GO:0008415; F:acyltransferase activity; IEA.
DR      GO; GO:0016740; F:transferase activity; IEA.
DR      InterPro; IPR003480; Transferase.
DR      Pfam; PF02458; Transferase; 1.
KW      Acyltransferase; Transferase.
SQ      SEQUENCE 469 AA; 48770 MW; 08C7B6076576A39C CRC64;

Query Match      100.0%; Score 32; DB 2; Length 469;
Best Local Similarity 100.0%; Pred.No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CRGDD 5
Db      201 CRGDD 205

RESULT 35
Q7PPD6
ID      Q7PPD6      PRELIMINARY;      PRT;      491 AA.
AC      Q7PPD6;
DT      01-MAR-2004 (TrEMBLrel. 26, Created)
DT      01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      ENSANGP0000021570 (Fragment).
GN      Name=ENSANGG00000019081;
OS      Anopheles gambiae str. PEST.
OC      Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC      Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX      NCBI_TaxID=180454;

```


RP SEQUENCE FROM N.A.
RA STRAIN=PEST;
RL Anopheles Genome Sequencing Consortium;
RR Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
RS -1- CAUTION: The sequence shown here is derived from an
SC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AAB01008952; EAA10613.2; -.
DR InterPro; IPR005123; 2OG-FeII_Oxy.
DR InterPro; IPR001440; TPR.
DR InterPro; IPR008941; TPR-like.
DR Pfam; PF03171; 2OG-FeII_Oxy; 1.
DR PROSITE; PSS0005; TPR; 1.
DR PROSITE; PSS0293; TPR_REGION; 1.
FT NON TER 1
SQ SEQUENCE 491 AA; 56112 MW; DCB3F52923F03BE6 CRC64;

Query Match 100.0%; Score 32; DB 2; Length 491;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
Db 252 CRGDD 256

RESULT 36
Q859H6 PRELIMINARY; PRT; 497 AA.
AC Q859H6;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
OS Bacteriophage f237.
OC Viruses.
OX NCBI_TaxID=221993;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22562669; PubMed=12676700;
RX DOI=10.1128/AEM.69.4.2194-2200.2003;
RA Myers M.L., Panicker G., Bej A.K.;
RT "PCR detection of a newly emerged pandemic Vibrio parahaemolyticus
RT O3:k6 pathogen in pure cultures and seeded waters from the Gulf of
RT Mexico.";
RL Appl. Environ. Microbiol. 69:2194-2200(2003).
DR EMBL; AY196694; AAO46790.1; -.
KW Hypothetical protein.
SQ SEQUENCE 497 AA; 55709 MW; 13EB96600D557B74 CRC64;

Query Match 100.0%; Score 32; DB 2; Length 497;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
Db 97 CRGDD 101

RESULT 37
Q9MCC5 PRELIMINARY; PRT; 497 AA.
AC Q9MCC5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Vpf497.
GN Names=vpf497;
OS Bacteriophage VFO3K6.
OC Viruses; ssDNA viruses; Inoviridae; Inovirus.
OX NCBI_TaxID=127507;
RN [1]

RP SEQUENCE FROM N.A.
RA Chang B., Miyamoto H., Ogawa M., Yoshida S., Taniguchi H.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB043678; BAA96428.1; -.
SQ SEQUENCE 497 AA; 55788 MW; 274E868DEB6E8B19 CRC64;

Query Match 100.0%; Score 32; DB 2; Length 497;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
Db 97 CRGDD 101

RESULT 38
Q9KGQ6 PRELIMINARY; PRT; 497 AA.
AC Q9KGQ6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein orf8 (Bacteriophage f237 ORF8).
GN Name=orf8; OrderedLocustNames=VP1561;
OS Vibrio parahaemolyticus.
OG Plasmid pO3K6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KKV237; PLASMID=PO3K6;
RX MEDLINE=20295086; PubMed=10834969;
RA Nasu H., Iida T., Sugahara T., Yamaichi Y., Park K., Yokoyama K.,
RA Makino K., Shinagawa H., Honda T.;
RT "A filamentous phage associated with recent pandemic Vibrio
RT parahaemolyticus O3:k6 strains.";
RL J. Clin. Microbiol. 38:2156-2161(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=RIMD 2210633 / Serotype O3:k6;
RX MEDLINE=22508454; PubMed=12620739; DOI=10.1016/S0140-6736(03)12659-1;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae.";
RL Lancet 361:743-749(2003).
DR EMBL; AP000581; BAA97266.1; -.
DR EMBL; AP005078; BAC59824.1; -.
KW Complete proteome; Hypothetical protein; Plasmid.
SQ SEQUENCE 497 AA; 55788 MW; 274E868DEB6E8B19 CRC64;

Query Match 100.0%; Score 32; DB 2; Length 497;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
Db 97 CRGDD 101

RESULT 39
Q8QS61 PRELIMINARY; PRT; 640 AA.
AC Q8QS61;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE UL27.
OS Pongine herpesvirus 4.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.

```

OX NCBI_TaxID=188763;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22421467; PubMed=12533697; DOI=10.1099/vir.0.18606-0;
RA Davidson A.J., Dolan A., Akter P., Addison C., Dargan D.J.,
RA Alencor D.J., McGeoch D.J., Hayward G.S.;
RT "The human cytomegalovirus genome revisited: comparison with the
RL chimpanzee cytomegalovirus genome.";
RL J. Gen. Virol. 84:17-28(2003).
DR EMBL; AF480884; AM00678.1; -
DR InterPro; IPR010302; Herpes_US.
DR Pfam; PF05999; Herpes_U5; 1.
SQ SEQUENCE 640 AA; 73457 MW; E4969DFC8BF4F34F CRC64;

Query Match 100.0%; Score 32; DB 2; Length 640;
Best Local Similarity 100.0%; Pred. No. 2.5e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0;

QY 1 CRGDD 5
DB 46 CRGDD 50

RESULT 40
Q73RY0
ID Q73RY0 PRELIMINARY; PRT; 776 AA.
AC Q73RY0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=MAP4298c;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AS017242; AA06848.1; -
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR003673; CAIB BAIF.
DR Pfam; PF02515; CoA_transf_3; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 776 AA; 82189 MW; 38A4528FC5B0B4BB CRC64;

Query Match 100.0%; Score 32; DB 2; Length 776;
Best Local Similarity 100.0%; Pred. No. 3.1e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0;

QY 1 CRGDD 5
DB 617 CRGDD 621

RESULT 41
Q85ZX4
ID Q85ZX4 PRELIMINARY; PRT; 934 AA.
AC Q85ZX4;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE LP03809P.
GN Name=Nrx-1;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Paclet J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celnik S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AV069731; AAL39876.1; -
DR HSSP; O63373; 1C4R.
DR FlyBase; FBgn0038975; Nrx-1.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR008985; Coa_like_lec_gl.
DR InterPro; IPR00742; EGF_2.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR001791; Laminin_G.
DR InterPro; IPR003129; TSP_N.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF02210; Laminin_G_2; 2.
DR SMART; SM00282; LamG; 2.
DR PROSITE; PS50026; EGF_3; 1.
DR PROSITE; PS50025; Lam_G_DOMAIN; 2.
SQ SEQUENCE 934 AA; 99518 MW; BC805CB3522A2EF8 CRC64;

Query Match 100.0%; Score 32; DB 2; Length 934;
Best Local Similarity 100.0%; Pred. No. 3.7e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0;

QY 1 CRGDD 5
DB 510 CRGDD 514

RESULT 42
MKH1_SCHPO
ID MKH1_SCHPO STANDARD; PRT; 1116 AA.
AC Q10407;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE MAP kinase kinase kinase mkhl (EC 2.7.1.-).
GN Name=mkhl; ORFNames=SPAC1F3.02c;
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SP66;
RX MEDLINE=97342589; PubMed=9199286;
RA Sengar A.S., Markley N.A., Marini N.J., Young D.;
RT "Mkhl, a MEK kinase required for cell wall integrity and proper
RT response to osmotic and temperature stress in Schizosaccharomycetes
RT pombe.";
RL Mol. Cell. Biol. 17:3508-3519(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360; DOI=10.1038/nature724;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,

```

RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Furnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of *Schizosaccharomyces pombe*.";
RL Nature 415:871-880(2002).
CC -!- FUNCTION: May regulate cell morphology, cell wall integrity, salt
CC resistance, cell cycle reentry from stationary-phase arrest, and
CC filamentous growth in response to stress. Activates the MAP kinase
CC kinase skh1/pekl by phosphorylation.
CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. MAP
CC kinase kinase subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U53872; AAB62319.1; -;
CC EMBL; 270690; CAA94620.1; -;
CC FIR; T38073; T38073.
CC HSP; Q16539; IKV1.
CC GeneDB_Spombe; SPAC1F3.02c; -;
CC InterPro; IPR011009; Kinase like.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR001660; SAM_kinase.
CC InterPro; IPR010993; SAM_homology.
CC InterPro; IPR008271; Ser_thr_pkin_AS.
CC InterPro; IPR002290; Ser_thr_pkinase.
CC Pfam; PF00069; Pkinase; 1.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00220; S_TKc; 1.
CC SMART; SM00454; SAM; 1.
CC PROSITE; PS00107; PROTEIN KINASE_ATP; 1.
CC PROSITE; PS00111; PROTEIN KINASE_DOM; 1.
CC PROSITE; PS00108; PROTEIN KINASE_ST; 1.
CC ATP-binding; Serine/threonine-protein kinase; Transferase.
KW DOMAIN 825 1094 Protein kinase.
FT NP_BIND 831 839 ATP (By similarity).
FT BINDING 854 854 ATP (By similarity).
FT ACT_SITE 955 955 Proton acceptor (By similarity).
SQ SEQUENCE 1116 AA; 125132 MW; 7AFDB3EC62EED47B CRC64;
Query Match 100.0%; Score 32; DB 1; Length 1116;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CRGDD 5
Db 677 CRGDD 681
RESULT 43
Q8JKT0 PRELIMINARY; PRT; 1145 AA.
AC Q8JKT0;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Orf31.
DE Name=orf31;
GN Heliothis zea virus 1.
OS Viruses; dsDNA viruses, no RNA stage; unclassified dsDNA viruses.

OX NCBI_TaxID=29250;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22174892; PubMed=12186886;
RX DOI=10.1128/JVI.76.18.9024-9034.2002;
RA Cheng C.H., Liu S.M., Chow T.Y., Hsiao Y.Y., Wang D.P., Huang J.J.,
RA Chen H.H.;
RT "Analysis of the complete genome sequence of the H2-1 virus suggests
RT that it is related to members of the Baculoviridae.";
RL J. Virol. 76:9024-9034(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Cheng C.H., Liu H.M., Hsiao Y.Y., Chow T.Y., Chen H.H.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Chao Y.;
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Chen H.H., Yeh W.B., Tso D.J.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Cheng C.H., Hsiao Y.Y., Liu S.M., Chow T.Y., Chen H.H.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF451898; AAN04326.1; -;
SQ SEQUENCE 1145 AA; 125435 MW; 80DCB38DCDC7DC42 CRC64;
Query Match 100.0%; Score 32; DB 2; Length 1145;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CRGDD 5
Db 17 CRGDD 21
RESULT 44
Q7PSM9 PRELIMINARY; PRT; 1504 AA.
AC Q7PSM9;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000010133 (Fragment).
GN Name=ENSANGG0000007644;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AAB801008817; EAA05336.2; -;
DR HSP; Q63373; IC4R.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR008985; Cona_like_lec_gl.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR001791; Laminin_G.
DR InterPro; IPR003129; TSP_N.
DR Pfam; PF02210; Laminin_G_2; 5.
DR Pfam; PF02210; EGF_3; 3.
DR PROSITE; PS50026; EGF_3; 3.
DR PROSITE; PS50025; LAM_G_DOMAIN; 4.
FT NON_TER 1

RA	SVIRSKAS R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C., Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.; "Finishing a whole-genome shotgun: Release 3 of the Drosophila melanogaster euchromatic genome sequence."; Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002). [3]
RX	SEQUENCE FROM N.A. MEDLINE=22426070; PubMed=12537573;
RA	Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirkas R., Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M., Ashburner M., Celniker S.E.; "The transposable elements of the Drosophila melanogaster euchromatin: a genomics perspective."; Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002). [4]
RN	SEQUENCE FROM N.A. MEDLINE=22426069; PubMed=12537572;
RP	Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S., Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E., Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P., Barrington R.C., Rogers Y.H., Blazek R.G., Champagne M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L., Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D., Ballou R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S., Borkova D., Botchan M.R., Bouck J., Brokstein P., Brattier P., Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W., Foster K., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K., Glocke A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Hariss N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C., Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Milehina N.V., Mobarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphree L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nusaker D.R., Pacle J.M., Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.Y., Wasserman D.A., Weinstock G.K., Weissbach J., Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.F., Zaveri J.S., Zhao M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.; "The genome sequence of Drosophila melanogaster."; Science 287:2185-2195(2000). [2]
RQ	QUERY MATCH Best Local Similarity 100.0%; Score 32; DB 2; Length 1837; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 CRGDD 5
DB	1413 CRGDD 1417
RESULT 46	
Q7QP19	PRELIMINARY; PRT; 73 AA.
ID	Q7QP19
AC	Q7QP19
DT	01-MAR-2004 (TrEMBLrel. 26, Created)
DT	01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE	GLP_83_1716_1495.
OS	Giardia lamblia ATCC 50803.
OC	Eukaryota; Diplomonadida; Hexamitidae; Giardiinae; Giardia.

RA	SVIRSKAS R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C., Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.; "Finishing a whole-genome shotgun: Release 3 of the Drosophila melanogaster euchromatic genome sequence.";
RT	Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN	[3]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=22426070; PubMed=12537573;
RA	Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskaas R., Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M., Ashburner M., Celniker S.E.;
RA	"The transposable elements of the Drosophila melanogaster euchromatin: a genomics perspective.";
RT	Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN	[4]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=22426069; PubMed=12537572;
RA	Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S., Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E., Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P., Barriscourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A., Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q., Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M., Lewis S.E.;
RA	"Annotation of the Drosophila melanogaster euchromatic genome: a systematic review.";
RT	Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN	[5]
RP	SEQUENCE FROM N.A.
RX	FlyBase;
RG	Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RL	[6]
RP	SEQUENCE FROM N.A.
RX	FlyBase;
RL	Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AE003739; AAF56006.1; -.
DR	HSSP; Q63373; IC4R.
DR	FlyBase; FBGN0038975; NrX-1.
DR	GO; GO:0005198; F:structural molecule activity; IEA.
DR	GO; GO:0007155; P:cell adhesion; IEA.
DR	InterPro; IPR000985; ConA_like_lec_gl.
DR	InterPro; IPR000742; EGF_2.
DR	InterPro; IPR006209; EGF-like.
DR	InterPro; IPR006210; IEGF.
DR	InterPro; IPR001791; Laminin_G.
DR	InterPro; IPR003129; TSP_N.
DR	Pfam; PF00008; EGF; 3.
DR	Pfam; PF02210; Laminin_G_2; 6.
DR	SMART; SM00181; EGF; 2.
DR	SMART; SM00282; LamG; 6.
DR	PROSITE; PS50026; EGF_3; 3.
DR	PROSITE; PS50025; LAM_G_DOMAIN; 6.
SQ	SEQUENCE 1837 AA; 19296 MW; 58D354C1CE7B6687 CRC64;
Query Match	
Best Local Similarity 100.0%; Score 32; DB 2; Length 1837;	
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 CRGDD 5
Db	1413 CRGDD 1417
RESULT 46	
Q7QP19	PRELIMINARY; PRT; 73 AA.
ID	AC Q7QP19;
DT	01-MAR-2004 (TrEMBLrel. 26; Created)
DT	01-MAR-2004 (TrEMBLrel. 26; Last sequence update)
DE	GLP_83_1716_1495.
OS	Giardia lamblia ATCC 50803.
OC	Eukaryota; Diplomonadida; Hexamitidae; Giardiinae; Giardia.
SQL	SEQUENCE 1504 AA; 163165 MW; 888039D2F5246488 CRC64;
Query Match	
Best Local Similarity 100.0%; Score 32; DB 2; Length 1504;	
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 CRGDD 5
Db	1085 CRGDD 1089
RESULT 45	
Q9VCZ9	PRELIMINARY; PRT; 1837 AA.
ID	AC Q9VCZ9;
DT	01-MAY-2000 (TrEMBLrel. 13; Created)
DT	01-MAY-2000 (TrEMBLrel. 13; Last sequence update)
DT	01-MAR-2004 (TrEMBLrel. 26; Last annotation update)
DE	C707050-PA.
DN	Names-Nrx-1; ORFNames-CG7050;
OS	Drosophila melanogaster (Fruit fly).
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC	Ephoroidea; Drosophilidae; Drosophila.
OX	NCBI_TaxID=7227;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L., Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D., Ballou R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S., Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P., Burtis K.C., Busan M.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W., Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K., Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Hariss N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C., Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Lascko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Milehina N.V., Mobarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nusaker D.R., Pacle J.M., Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.Y., Watersman D.A., Weinstock G.K., Weissbach J., Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.F., Zaveri J.S., Zhao M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;"
RT	"The genome sequence of Drosophila melanogaster.";
RL	Science 287:2185-2195 (2000).
RN	[2]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=22426065; PubMed=12537568;
RA	Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A., Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A., George R.A., Park S., Pfeiffer B.D., Richards S., Sodergren E.J., Pacle J.M., Harkins R., Lavery T., Muzny D.M., Nelson C.R.,

OX NCBI_TaxID=184922;
 RN SEQUENCE FROM N.A.
 RC STRAIN=WB C6;
 RA Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D.,
 RA Olsen G.J., Sogin M.L.;
 RT "Draft sequence of the Giardia lamblia genome.";
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AACB0100181; EAA36741.1; -;
 SQ SEQUENCE 73 AA; 8125 MW; 4F866C8AEF254DDE CRC64;

Query Match 90.6%; Score 29; DB 2; Length 73;

Best Local Similarity 80.0%; Pred. No. 1.3e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5

Db 30 CRGDD 34

RESULT 47

Q69789 ID Q69789 PRELIMINARY; PRT; 112 AA.

AC Q69789;

DT 25-OCT-2004 (TrEMBLrel. 28, Created)

DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE Hypothetical protein P0652D10.11.

GN Name=P0652D10.11;

OS Oryza sativa (japonica cultivar-group).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzaceae; Oryza.

OC NCBI_TaxID=39947;

OX [1]

RN SEQUENCE FROM N.A.

RA Sasaki T., Matsumoto T., Yamamoto K.;

RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC

RT clone:P0652D10.";

RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AP004757; BAD33235.1; -;

KW Hypothetical protein.

SQ SEQUENCE 112 AA; 12420 MW; 8BEA619585F664F9 CRC64;

Query Match

Best Local Similarity 90.6%; Score 29; DB 2; Length 112;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5

Db 45 CRGDD 49

RESULT 48

Q726G7

ID Q726G7 PRELIMINARY; PRT; 126 AA.

AC Q726G7;

DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Lipoprotein, putative.

GN OrderedLocNames=DVU3141;

OS Desulfovibrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB

OS 8303).

OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales;

OC Desulfovibrionaceae; Desulfovibrio.

OC NCBI_TaxID=882;

OX [1]

RN SEQUENCE FROM N.A.

RP PubMed=15077118; DOI=10.1038/nbt959;

RX

RA Heidelberg J.F., Seshadri R., Haveman S.A., Hemme C.L., Paulsen I.T.,
 RA Kolonay J.F., Eisen J.A., Ward N.D., Meche B.A., Brinkac L.M.,
 RA Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S., Madupu R.,
 RA Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,
 RA Peterson J.D., Davidson T.M., Zafar N., Zhou L., Radune D.,
 RA Dimitrov G., Hance M., Tran K., Khouri H.M., Gill J., Utterback T.R.,
 RA Feldblyum T.V., Wall J.D., Voordouw G., Fraser C.M.;
 RT "The genome sequence of the anaerobic, sulfate-reducing bacterium
 RT Desulfovibrio vulgaris Hildenborough.";
 RL Nat. Biotechnol. 22:554-559(2004).
 DR EMBL; AE017319; AAS97611.1; -;
 DR TIGR; DVU3141; -;
 KW Complete proteome; Lipoprotein.
 SQ SEQUENCE 126 AA; 13270 MW; DB78BD19E539D9EB CRC64;

Query Match

Best Local Similarity 90.6%; Score 29; DB 2; Length 126;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5

Db 106 CRGDD 110

RESULT 49

Q8XNS3

ID Q8XNS3 PRELIMINARY; PRT; 168 AA.

AC Q8XNS3;

DT 01-MAR-2002 (TrEMBLrel. 20, Created)

DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Probable transcriptional regulator.

OS OrderedLocNames=CPE0259;

OC Clostridium perfringens.

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

OC Clostridium.

OC NCBI_TaxID=1502;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=13;

RX MEDLINE=21664373; PubMed=11792842; DOI=10.1073/pnas.022493799;

RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamaehita A.,

RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;

RT "Complete genome sequence of Clostridium perfringens, an anaerobic

RT flesh-eater.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).

DR EMBL; AP003186; BAB79965.1; -;

DR GO; GO:0003700; P:transcription factor activity; IEA.

DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.

DR Pfam; PF02559; Card_TRCF; 1.

KW Complete proteome.

SQ SEQUENCE 168 AA; 19582 MW; B6AC5EAA335AF717 CRC64;

Query Match

Best Local Similarity 90.6%; Score 29; DB 2; Length 168;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5

Db 129 CRGDD 133

RESULT 50

Q8T399

ID Q8T399 PRELIMINARY; PRT; 225 AA.

AC Q8T399;

DT 01-JUN-2002 (TrEMBLrel. 21, Created)

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Putative coagulation serine protease.

GN Name=sp5;

OS Ciona intestinalis.

OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;

OC Phlebobranchia; Cionidae; Ciona.
OX NCBI_TaxID=7719;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Repatopancreas;
RA Hammond J.A., Nakao M., Yano T., Kemp G.D., Smith V.J.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; AJ431687; CAD24310.1; -
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR009003; Pept Ser_Cys.
DR PRINTS; PR00722; CHYMOTRFSIN.
DR SMART; SM00020; TRYPSIN_DOM; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
KW Hydrolyase; Protease; Serine protease.
SQ SEQUENCE 225 AA; 24146 MW; 0EC9AE25D7CB743B CRC64;

Query Match 90.6%; Score 29; DB 2; Length 225;
Best Local Similarity 80.0%; Pred. No. 3.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
Db 193 CRGDD 197
|:|:|

RESULT 51
Q9NSJ7 PRELIMINARY; PRT; 264 AA.
AC Q9NSJ7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein DKFZp434K1822 (Fragment).
GN Name=DKFZp434K1822;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Blum H., Bauersachs S., Mewes H.W., Weil B., Wiemann S.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL162082; CAB82415.1; -
DR PIR; T47183; T47183.
DR GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.
DR GO; GO:0004221; F:ubiquitin thiolesterase activity; IEA.
DR GO; GO:0006511; P:ubiquitin-dependent protein catabolism; IEA.
DR InterPro; IPR000345; Cyt c heme_BS.
DR InterPro; IPR001394; Peptidase_C19.
DR Pfam; PF00443; UCH; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_2.
DR PROSITE; PS00973; UCH_2_2; 1.
DR PROSITE; PS02035; UCH_2_3; 1.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 264 AA; 29783 MW; 60D25351F7BE8464 CRC64;

Query Match 90.6%; Score 29; DB 2; Length 264;
Best Local Similarity 80.0%; Pred. No. 4.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
Db 17 CRGDD 21
|:|:|

RESULT 52
Q8CUR7 PRELIMINARY; PRT; 276 AA.
AC Q8CUR7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Phosphomethylpyrimidine kinase.
GN Name=thiD; OrderedLocusNames=OB1040;
OS Oceanobacillus theyensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.
OX NCBI_TaxID=182710;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HT881;
RX MEDLINE=22220767; PubMed=12235376; DOI=10.1093/nar/gkf526;
RA Takami H., Takaki Y., Uchiyama I.;
RT "Genome sequence of Oceanobacillus theyensis isolated from the Iheya
RT Ridge and its unexpected adaptive capabilities to extreme
RT environments";
RL Nucleic Acids Res. 30:3927-3935(2002).
DR EMBL; AP004596; BAC12996.1; -
DR HSSP; P55882; IUXH.
DR GO; GO:0008972; F:phosphomethylpyrimidine kinase activity; IEA.
DR GO; GO:0009228; P:thiamin biosynthesis; IEA.
DR InterPro; IPR004399; HMP-P_kinase.
DR TIGRFAMs; TIGR00097; HMP-P_kinase; 1.
KW Complete proteome.
SQ SEQUENCE 276 AA; 29916 MW; 65D5E6A236F216B9 CRC64;

Query Match 90.6%; Score 29; DB 2; Length 276;
Best Local Similarity 80.0%; Pred. No. 4.8e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
Db 110 CRGDD 114
|:|:|

RESULT 53
Q92XS2 PRELIMINARY; PRT; 296 AA.
AC Q92XS2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN ORFNames=Smaz147;
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Plasmid pSymA.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396509; PubMed=11481432; DOI=10.1073/pnas.161294798;
RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
RA Barlett-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,
RA Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,
RA Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;
RT "Nucleotide sequence and predicted functions of the entire
RT Sinorhizobium meliloti pSymA megaplasmid";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
DR EMBL; AE007302; AAK65827.1; -
DR PIR; A95408; A95408.
DR InterPro; IPR011111; Plasmid_RepB.
DR Pfam; PF07506; RepB; 1.
KW Complete proteome; Hypothetical protein; Plasmid.
SQ SEQUENCE 296 AA; 32439 MW; D7D0AF683E0AE178 CRC64;

Query Match 90.6%; Score 29; DB 2; Length 296;

Best Local Similarity 80.0%; Pred. No. 5.1e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 1;

QY 1 CRGDD 5
|:|:|

Db 5 CKGDD 9

RESULT 54
Q8BJG3 PRELIMINARY; PRT; 302 AA.

AC Q8BJG3
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE Mus musculus 12 days embryo spinal ganglion cDNA, RIKEN full-length
DE enriched library, clone: D130076P14 product: similar to UBIQUITIN
DE SPECIFIC PROTEASE (Fragment).
GN Name=Usp22;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal ganglion;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal ganglion;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal ganglion;
RX MEDLINE=20499374; PubMed=11042139; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal ganglion;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kiteunai T., Tachiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal ganglion;
RX Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,

RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Kori F., Inotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK084022; BAC39100.1; -;
DR MEROPS; C19_035; -;
DR MGD; MGI:2144157; Usp22.
DR GO; GO:0004197; F:peptidase-type endopeptidase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004221; F:ubiquitin thiolesterase activity; IEA.
DR GO; GO:0006511; P:ubiquitin-dependent protein catabolism; IEA.
DR InterPro; IPR000345; CytC heme BS.
DR InterPro; IPR001394; Peptidase_C19.
DR Pfam; PF00443; UCH; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_2.
DR PROSITE; PS00973; UCH_2_2; 1.
DR PROSITE; PS0235; UCH_2_3; 1.
KW Protease.
FT NON_TER
SQ SEQUENCE 302 AA; 34288 MW; 3E0A5CFA7CD8F72C CRC64;
Query Match 90.6%; Score 29; DB 2; Length 302;
Best Local Similarity 80.0%; Pred. No. 5.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
|:|:|

Db 55 CKGDD 59

RESULT 55
Q7MX77 PRELIMINARY; PRT; 319 AA.

AC Q7MX77
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Glycosyl transferase, group 2 family protein.
DE OrderedLocustNames=PG0334;
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroides (Class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=W83;
RX MEDLINE=22829867; PubMed=12949112;
RX DOI=10.1128/JB.185.18.5591-5601.2003;
RA Nelson K.E., Fleischmann R.D., DeBoy R.T., Paulsen I.T., Fouts D.E.,
RA Eisen J.A., Daugherty S.C., Dodson R.J., Durkin A.S., Gwinn M.L.,
RA Haft D.H., Kolonay J.F., Nelson W.C., Mason T.M., Tallon L., Gray J.,
RA Granger D., Tettelin H., Dong H., Galvin J.L., Duncan M.J.,
RA Dewhirst F.E., Fraser C.M.;
RT "Complete genome sequence of the oral pathogenic bacterium
RT Porphyromonas gingivalis strain W83.";
RL J. Bacteriol. 185:5591-5601(2003).
DR EMBL; AB017173; AAQ65547.1; -;
DR TIGR; PG0334; -;
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR001173; Glyco_trans_2.
DR Pfam; PF00535; Glycoe_transf_2; 1.
KW Complete proteome; Transferase.
SQ SEQUENCE 319 AA; 36313 MW; 73E349629E9F80CD CRC64;
Query Match 90.6%; Score 29; DB 2; Length 319;
Best Local Similarity 80.0%; Pred. No. 5.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 CRGDD 5
DB      63 CRGDD 67

RESULT 56
Q940P2  PRELIMINARY;      PRT;      324 AA.
AC      Q940P2;
DT      01-DEC-2001 (TrEMBLrel. 19, Created)
DT      01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT      05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE      Atlg72050/F28P5.6.
OS      Arabidopsis thaliana (Mouse-ear cress).
OC      Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC      eurosids I; Brassicales; Brassicaceae; Arabidopsie.
OX      NCBI_TaxID=3702;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
RA      Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,
RA      Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
RA      Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
RA      Miranda M., Narusaka M., Nguyen T., Onodera C.S., Palm C.J.,
RA      Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA      Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
RA      Shinozaki K., Davis R.W., Theologis A., Ecker J.R.,
RA      Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Kim C.J., Chen H., Cheuk R., Meyers M.C., Banh J., Bowser L.,
RA      Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA      Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA      Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen T., Onodera C.S.,
RA      Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA      Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,
RA      Shinozaki K., Davis R.W., Theologis A., Ecker J.R.,
RA      Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RL      EMBL; AY054225; AL06885.1; -
DR      EMBL; AY066042; AL47409.1; -
DR      GO; GO:0005634; C:nucleus; IEA.
DR      GO; GO:0003676; F:nucleic acid binding; IEA.
DR      GO; GO:0008270; F:zinc ion binding; IEA.
DR      InterPro; IPR007087; Znf_C2H2.
DR      Pfam; PF00096; zf-C2H2; 5.
DR      SMART; SM00355; Znf_C2H2; 7.
DR      PROSITE; PS00028; ZINC_FINGER_C2H2_1; 6.
DR      PROSITE; PS0157; ZINC_FINGER_C2H2_2; 4.
DR      SEQUENCE 324 AA; 36527 MW; ED4AB6747A2FCE34 CRC64;

Query Match      90.6%; Score 29; DB 2; Length 324;
Best Local Similarity 80.0%; Pred. No. 5.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 CRGDD 5
DB      81 CRGDD 85

RESULT 57
Q9LKV3  PRELIMINARY;      PRT;      360 AA.
AC      Q9LKV3;
DT      01-OCT-2000 (TrEMBLrel. 15, Created)
DT      01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Ptl1 kinase-like protein.
GN      Name=Ptl1b;
OS      Glycine max (Soybean).
OC      Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

```

```

OC      eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX      NCBI_TaxID=3847;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      MEDLINE=21036115; PubMed=11185558; DOI=10.1006/abbi.2000.2080;
RA      Staswick P.;
RT      "Two expressed soybean genes with high sequence identity to tomato
RT      Ptl1 kinase lack autophosphorylation activity.";
RL      Arch. Biochem. Biophys. 383:233-237(2000).
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Staswick P.E.;
RA      Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF249318; AAF91337.1; -
DR      HSSP; P00523; 2PTK.
DR      GO; GO:0005524; F:ATP binding; IEA.
DR      GO; GO:0004672; F:protein kinase activity; IEA.
DR      GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR      InterPro; IPR011009; Kinase like.
DR      InterPro; IPR00719; Prot Kinase.
DR      ProDom; PD000001; Prot Kinase; 1.
DR      PROSITE; PS00107; PROTEIN_KINASE_DOM; 1.
DR      PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW      ATP-binding; Kinase.
SQ      SEQUENCE 360 AA; 39481 MW; 0B54143D695C2B90 CRC64;

Query Match      90.6%; Score 29; DB 2; Length 360;
Best Local Similarity 80.0%; Pred. No. 6.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 CRGDD 5
DB      7 CRGDD 11

RESULT 58
Q9LKV4  PRELIMINARY;      PRT;      360 AA.
AC      Q9LKV4;
DT      01-OCT-2000 (TrEMBLrel. 15, Created)
DT      01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Ptl1 kinase-like protein.
GN      Name=Ptl1a;
OS      Glycine max (Soybean).
OC      Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC      eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX      NCBI_TaxID=3847;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      MEDLINE=21036115; PubMed=11185558; DOI=10.1006/abbi.2000.2080;
RA      Staswick P.;
RT      "Two expressed soybean genes with high sequence identity to tomato
RT      Ptl1 kinase lack autophosphorylation activity.";
RL      Arch. Biochem. Biophys. 383:233-237(2000).
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Staswick P.E.;
RA      Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF249317; AAF91336.1; -
DR      HSSP; P00523; 2PTK.
DR      GO; GO:0005524; F:ATP binding; IEA.
DR      GO; GO:0004672; F:protein kinase activity; IEA.
DR      GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR      InterPro; IPR011009; Kinase like.
DR      InterPro; IPR00719; Prot Kinase.
DR      ProDom; PD000001; Prot Kinase; 1.
DR      PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR      PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW      ATP-binding; Kinase.
SQ      SEQUENCE 360 AA; 39375 MW; E170D95DD20E2D6F CRC64;

```



```

Query Match          90.6%; Score 29; DB 2; Length 360;
Best Local Similarity 80.0%; Pred. No. 6.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
DB 7 CKGDD 11

RESULT 59
Q72MX9 PRELIMINARY; PRT; 366 AA.
AC Q72MX9;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE LipL36.
GN Names=lipL36; OrderedLocusNames=LIC13060;
OS Leptospira interrogans (serogroup Icterohaemorrhagiae / serovar Copenhageni).
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OC NCBI_TaxID=44275;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Fiocruz L1-130;
RX PubMed=15028702; DOI=10.1128/JB.186.7.2164-2172.2004;
RA Nascimento A.L.T.O., Ko A.I., Martins E.A.L., Monteiro-Vitorello C.B., Ho P.L., Haake D.A., Verjovski-Almeida S., Hartekeel R.A., Marques M.V., Oliveira M.C., Mencek C.F.M., Leite L.C.C., Carrer H., Coutinho L.L., Degraeve W.M., Dellagostin O.A., El-Dorry H., Ferro E.S., Ferro M.I.T., Furlan L.R., Gamberini M., Gigliotti E.A., Goes-Neto A., Goldman G.H., Goldman M.H.S., Harakava R., Jeronimo S.M.B., Junqueira-de-Azevedo J.L.M., Kimura E.T., Kuranai E.E., Lemos E.G.M., Lemos M.V.F., Marino C.L., Nunes L.R., de Oliveira R.C., Pereira G.G., Reis M.S., Schrieffer A., Siqueira W.J., Sommer P., Tsai S.M., Simpson A.J.G., Perro J.A., Camargo L.E.A., Kitajima J.P., Setubal J.C., Van Sluys M.A.;
RA "Comparative genomics of two Leptospira interrogans serovars reveals novel insights into physiology and pathogenesis.";
RL J. Bacteriol. 186:2164-2172(2004).
DR EMBL; AB017299; AA571609.1; -.
KW Complete proteome.
SQ SEQUENCE 366 AA; 37620 MW; 68924462DADA1F73 CRC64;

Query Match          90.6%; Score 29; DB 2; Length 366;
Best Local Similarity 80.0%; Pred. No. 6.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
DB 21 CKGDD 25

RESULT 60
Q8F8R3 PRELIMINARY; PRT; 366 AA.
AC Q8F8R3;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE LipL36 protein.
GN Names=lipL36; OrderedLocusNames=LA0492;
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OC NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RX MEDLINE=22598143; PubMed=12712204; DOI=10.1038/nature01597;
RA Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H., Zhang Y.-X., Xiong H., Lu G., Lu L.-F., Jiang H.-Q., Jia J., Tu Y.-F., Jiang J.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-P., Zhang Y., Zhu G.-F., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,

```

```

RA Yao Z.-J., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Danchin A., Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z., Xu J.-G., Zhao G.-P.;
RA "Unique physiological and pathogenic features of Leptospira RT interrogans revealed by whole-genome sequencing.";
RL Nature 422:888-893(2003).
DR EMBL; AB011235; AAN47691.1; -.
KW Complete proteome.
SQ SEQUENCE 366 AA; 37677 MW; 106A18B837388EC5 CRC64;

Query Match          90.6%; Score 29; DB 2; Length 366;
Best Local Similarity 80.0%; Pred. No. 6.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
DB 21 CKGDD 25

RESULT 61
Q6FIV5 PRELIMINARY; PRT; 386 AA.
AC Q6FIV5;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Similar to sp|Q12385 Saccharomyces cerevisiae YLR099c ICT1. GN ORFNames=CAGL0M11462g;
OS Candida glabrata CBS138.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=284593;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS138;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S., Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E., Goffard N., Frangoul L., Aigle M., Anthouard V., Babour A., Barbe V., Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C., Boisarame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A., Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A., Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R., Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H., Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O., Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A., Swennene D., Tekia F., Wesolowski-Louvel M., Westhof E., Wirth B., Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A., Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J., Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
DR EMBL; CR380959; CAG62819.1; -.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0006725; P:aromatic compound metabolism; IEA.
DR InterPro; IPR000073; A/B hydrolase.
DR InterPro; IPR003089; AB hydrolase.
DR InterPro; IPR003079; Ser esters.
DR Pfam; PF00561; Abhydrolase 1; 1.
DR PRINTS; PR00111; ABHYDROLASE.
SQ SEQUENCE 386 AA; 44552 MW; 1F58363365217E51 CRC64;

Query Match          90.6%; Score 29; DB 2; Length 386;
Best Local Similarity 80.0%; Pred. No. 6.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
DB 3 CKGDD 7

RESULT 62
Q8EV52

```

ID Q6EV52 PRELIMINARY; PRT; 403 AA.
 AC Q6EV52;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DE Protein-O-fucosyltransferase 2 (Fragment).
 GN Name=fut13;
 OS Ciona savignyi.
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 OC Phlebobranchia; Clonidae; Ciona.
 OX NCBI_TaxID=51511;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=12966037; DOI=10.1093/glycob/cwgl13;
 RA Martinez-Duncker I., Mollicone R., Candelier J.J., Breton C.,
 RA Oriol R.;
 RT "A new superfamily of protein-O-fucosyltransferases, alpha2-
 RT fucosyltransferases and alpha6-fucosyltransferases: phylogeny and
 RT identification of conserved peptide motifs.";
 RL Glycobiology 13:1C-5C(2003).
 [2]
 RP SEQUENCE FROM N.A.
 RA Martinez-Duncker I., Oriol R., Mollicone R.;
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ781760; CAG03735.1; -;
 DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
 KW Glycosyltransferase; Transferase.
 FT NON_TER 1
 SQ SEQUENCE 403 AA; 47347 MW; E930EB3E3A23C416 CRC64;

 Query Match 90.6%; Score 29; DB 2; Length 403;
 Best Local Similarity 80.0%; Pred. No. 6.9e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

 QY 1 CRGDD 5
 Db 38 CRGDD 42

 RESULT 63
 Q84MZ4 PRELIMINARY; PRT; 412 AA.
 ID Q84MZ4;
 AC Q84MZ4;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Transcription factor IIIA (Hypothetical protein) (Atlg72050).
 GN Name=TFIIIA; ORFNames=Atlg72050;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22598218; PubMed=12711688; DOI=10.1093/nar/gkg335;
 RA Mathieu O., Yukawa Y., Prieto J.L., Vaillant I., Sugiura M.,
 RA Tourmente S.;
 RT "Identification and characterization of transcription factor IIIA and
 RT ribosomal protein L5 from Arabidopsis thaliana.";
 RL Nucleic Acids Res. 31:2424-2433(2003).
 [2]
 RP SEQUENCE FROM N.A.
 RA Gong W., Shen Y.P., Ma L.G., Du Y.L., Wang D.H., Yang J.Y., Hu L.D.,
 RA Liu X.F., Dong C.X., Ma L., Chen Y.H., Yang X.Y., Gao Y., Zhu D.,
 RA Tan X., Mu J.Y., Zhang D.B., Liu Y.T., Dinesh-Kumar S.P., Li Y.,
 RA Wang X.P., Gu H.Y., Qu L.J., Bai S.N., Lu Y.T., Li J.Y., Zhao J.D.,
 RA Zuo J., Huang H., Deng X.W., Zhu Y.X.;
 RT "Genome-Wide ORFeome Cloning and Analysis of Arabidopsis Transcription
 RT Factor Genes.";
 RL Plant Physiol. 135:773-782(2004).
 [3]
 RP SEQUENCE FROM N.A.

RA Zhu Y.;
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Gong W., Pan Y., Du Y.-L., Wang D.-H., Yang J.-Y., Hu L.-D.,
 RA Dong C.-X., Ma L., Ma L.-G., Mu J.-Y., Zhang D.-B., Li Y., Wang X.-P.,
 RA Gu H.-Y., Qu L.-J., Bai S.-N., Lu Y.-T., Li J.-Y., Zhao J.-D.,
 RA Zuo J.-R., Huang H., Deng X.-W., Zhu Y.-X.;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY186610; AAO73339.1; -;
 DR EMBL; AJ630478; CAG25851.1; -;
 DR EMBL; AY568650; AAS79540.1; -;
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR GO; GO:0008270; F:zinc ion binding; IEA.
 DR InterPro; IPR007087; Znf_C2H2.
 DR Pfam; PF00096; zf-C2H2; 7.
 DR SMART; SM00355; Znf_C2H2; 9.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 7.
 DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 5.
 KW Hypothetical protein.
 SQ SEQUENCE 412 AA; 46653 MW; AC171P610724C09F CRC64;

 Query Match 90.6%; Score 29; DB 2; Length 412;
 Best Local Similarity 80.0%; Pred. No. 7.1e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

 QY 1 CRGDD 5
 Db 169 CKGDD 173

 RESULT 64
 Q9C7G8 PRELIMINARY; PRT; 413 AA.
 ID Q9C7G8;
 AC Q9C7G8;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE C2H2-type zinc finger protein, putative.
 GN Name=F28PS.6;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
 RA Maiti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R.,
 RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC069273; AAG51140.1; -;
 DR PIR; F96743; F96743.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR GO; GO:0008270; F:zinc ion binding; IEA.
 DR InterPro; IPR007087; Znf_C2H2.
 DR Pfam; PF00096; zf-C2H2; 6.
 DR SMART; SM00355; Znf_C2H2; 9.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 6.
 DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 5.
 SQ SEQUENCE 413 AA; 46592 MW; B9F505F90B4C94FD CRC64;

 Query Match 90.6%; Score 29; DB 2; Length 413;
 Best Local Similarity 80.0%; Pred. No. 7.1e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

 QY 1 CRGDD 5
 Db 191 CKGDD 195

```
RESULT 65
Q6MK81
ID Q6MK81 PRELIMINARY; PRT; 419 AA.
AC Q6MK81;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein precursor.
GN OrderedlocusNames=Bd529;
OS Bdellovibrio bacteriovorus.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Bdellovibrionales;
OC Bdellovibrionaceae; Bdellovibrio.
OX NCBI_TaxID=959;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HD100 / DSM 50701 / ATCC 15356 / NCIB 9529;
RX PubMed=14752164; DOI=10.1126/science.1093027;
RA Rendulic S., Jagtap P., Rosinus A., Eppinger M., Baar C., Lanz C.,
RA Keller H., Lambert C., Evans K.J., Gosemann A., Meyer F.,
RA Sockett R.E., Schuster S.C.;
RT "A predator unmasked: life cycle of Bdellovibrio bacteriovorus from a
RT genomic perspective.";
RL Science 303:689-692(2004).
DR EMBL; BX842653; CAE80328.1; -.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR009056; Cytochrome_c.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_2.
KW Complete proteome.
FT SIGNAL 1 10 Potential.
SQ SEQUENCE 419 AA; 44302 MW; B45746CAFAA17E2 CRC64;

Query Match 90.6%; Score 29; DB 2; Length 419;
Best Local Similarity 80.0%; Pred. No. 7.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
DB 163 CRGDD 167

RESULT 66
Q7YZS7
ID Q7YZS7 PRELIMINARY; PRT; 434 AA.
AC Q7YZS7;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein fut13.
GN Name=fut13;
OS Ciona intestinalis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC Phlebobranchia; Clonidae; Ciona.
OX NCBI_TaxID=7719;
RN [1]
RP SEQUENCE FROM N.A.
RA Martinez-Dunker I., Mollicone R., Candellier J.J., Breton C., Oriol R.;
RT "A new superfamily of protein-O-fucosyltransferases, alpha2-
RT fucosyltransferases, and alpha6-fucosyltransferases: phylogeny and
RT identification of conserved peptide motifs.";
RL Glycobiology 13:1c-5c(2003).
DR EMBL; AJ575656; CAE02609.1; -.
KW Hypothetical protein.
SQ SEQUENCE 434 AA; 50605 MW; D853C746D6E896C CRC64;

Query Match 90.6%; Score 29; DB 2; Length 434;
Best Local Similarity 80.0%; Pred. No. 7.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
DB 69 CRGDD 73

RESULT 67
Q6MZY4
ID Q6MZY4 PRELIMINARY; PRT; 453 AA.
AC Q6MZY4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein DKFZp686P16150 (fragment).
GN Name=DKFZp686P16150;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human colon endothel primary cell culture;
RG The German Human CDNA Consortium;
RA Bloecker H., Boecher M., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BX640815; CAE45893.1; -.
DR GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.
DR GO; GO:0004221; F:ubiquitin thiolesterase activity; IEA.
DR GO; GO:0006511; F:ubiquitin-dependent protein catabolism; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001394; Peptidase_C19.
DR Pfam; PF00443; UCH; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_2.
DR PROSITE; PS00972; UCH_2_1; 1.
DR PROSITE; PS00973; UCH_2_2; 1.
DR PROSITE; PS50235; UCH_2_3; 1.
KW Hypothetical protein.
FT NON_TER 1 1
SQ SEQUENCE 453 AA; 51965 MW; 9A513E25242A420C CRC64;

Query Match 90.6%; Score 29; DB 2; Length 453;
Best Local Similarity 80.0%; Pred. No. 7.8e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
DB 206 CRGDD 210

RESULT 68
Q8CEG8
ID Q8CEG8 PRELIMINARY; PRT; 519 AA.
AC Q8CEG8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus 12 days embryo head cDNA, RIKEN full-length enriched
DE library, clone:3002003A21 product:ubiquitin specific protease 27, X
DE chromosome, full insert sequence.
GN Name=Usp27x;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
```

RNA SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
[4]
RNA SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=20499374; PubMed11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayata N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
[5]
RNA SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=20530913; PubMed11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kusunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
[6]
RNA SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito K., Saichon K., Sakai K., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Tova T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK028249; BAC25840.1; -;
DR MGD; MGI:1859645; Usp27x.
DR GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004221; F:ubiquitin thiolesterase activity; IEA.
DR GO; GO:0006511; F:ubiquitin-dependent protein catabolism; IEA.
DR InterPro; IPR000345; CytC heme-BS.
DR InterPro; IPR001394; Peptidase_C19.
DR InterPro; IPR001607; Znf_UBP.
DR Pfam; PF00443; UCH; 1.
DR Pfam; PF02148; zf-UBP; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE; PS00972; UCH_2_1; 1.
DR PROSITE; PS0235; UCH_2_3; 1.
KW Protease.
SQ SEQUENCE 519 AA; 58106 MW; DCC088BEC616AC86 CRC64;

Query Match 90.6%; Score 29; DB 2; Length 519;
Best Local Similarity 80.0%; Pred. No. 8.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
|:|:|
DB 301 CKGDD 305

RESULT 69

Q6GNI6
ID Q6GNI6; PRELIMINARY; PRT; 523 AA.
AC Q6GNI6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE MGC82781 protein.
GN Name=MGC82781;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=89355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=22388257; PubMed12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalski U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=22341132; PubMed12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Klein S., Gerhard D.S.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC073524; AAH73524.1; -;
DR GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.
DR GO; GO:0004221; F:ubiquitin thiolesterase activity; IEA.
DR GO; GO:0006511; F:ubiquitin-dependent protein catabolism; IEA.
DR InterPro; IPR001394; CytC heme BS.
DR InterPro; IPR001607; Znf_UBP.
DR Pfam; PF00443; UCH; 1.
DR Pfam; PF02148; zf-UBP; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_2.
DR PROSITE; PS00972; UCH_2_1; 1.
DR PROSITE; PS00973; UCH_2_2; 1.
DR PROSITE; PS0235; UCH_2_3; 1.
SQ SEQUENCE 523 AA; 60048 MW; 731A0AE1C2E67D61 CRC64;
Query Match 90.6%; Score 29; DB 2; Length 523;
Best Local Similarity 80.0%; Pred. No. 9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
|:|:|
DB 276 CKGDD 280

```

RESULT 70
Q66JV8
ID Q66JV8 PRELIMINARY; PRT; 525 AA.
AC Q66JV8;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Ubiquitin specific protease 22.
GN Names:Usp22;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RA Director MGC Project;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC080737; AAH80737.1; -.
DR InterPro; IPR001394; Peptidase_C19.
DR Pfam; PF00443; UCH_1.
DR PROSITE; PS00973; UCH_2_2; 1.
DR PROSITE; PS0235; UCH_2_3; 1.
DR Protease.
KW Protease.
SQ SEQUENCE 525 AA; 59954 MW; F884ED9B2FBF988 CRC64;

Query Match 90.6%; Score 29; DB 2; Length 525;
Best Local Similarity 80.0%; Pred. No. 9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
Db 278 CXGDD 282

RESULT 71
Q9JIG5
ID Q9JIG5 PRELIMINARY; PRT; 525 AA.
AC Q9JIG5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Ubiquitin specific protease (Fragment).
GN Names:Usp27x; Synonyms:Uspx;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

```

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX PubMed=10857745; DOI=10.1006/geno.2000.6173;
RA Means G.D., Toy D.Y., Baum P.R., Derry J.M.J.;
RT "A transcript map of a 2-Mb BAC contig in the proximal portion of the
RT mouse X chromosome and regional mapping of the scurfy mutation.";
RL Genomics 65:213-223(2000).
DR EMBL; AF229643; AAF66953.1; -.
DR MEROPS; C19.075; -.
DR MGD; MGI:1859645; Usp27x.
DR GO; GO:0004197; P-cysteine-type endopeptidase activity; IEA.
DR GO; GO:0008233; P-peptidase activity; IEA.
DR GO; GO:0004221; Fubiquitin thiolesterase activity; IEA.
DR GO; GO:0006511; Fubiquitin-dependent protein catabolism; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001394; Peptidase_C19.
DR InterPro; IPR001607; Znf_UBP.
DR Pfam; PF00443; UCH_1.
DR Pfam; PF02148; zf-UBP; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE; PS00972; UCH_2_1; 1.
DR PROSITE; PS00973; UCH_2_2; 1.
DR PROSITE; PS0235; UCH_2_3; 1.
DR Protease.
KW Protease.
FT NON_TER 1
SQ SEQUENCE 525 AA; 59670 MW; 71B7F18E25A2720B CRC64;

Query Match 90.6%; Score 29; DB 2; Length 525;
Best Local Similarity 80.0%; Pred. No. 9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
Db 267 CXGDD 271

RESULT 72
Q8BSW2
ID Q8BSW2 PRELIMINARY; PRT; 540 AA.
AC Q8BSW2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus adult male pituitary gland cDNA, RIKEN full-length
DE enriched library, clone:530405I05 product:ubiquitin specific protease
DE 27, X chromosome, full insert sequence. (Fragment).
GN Name:Usp27x;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Pituitary gland;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Pituitary gland;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Pituitary gland;
RX The FANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";

```

```

RL Nature 420:563-573 (2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Pituitary gland;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630 (2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Pituitary gland;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuina T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishimi T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fuyiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771 (2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Pituitary gland;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Kori F., Inotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Hori H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Nunazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akashira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AK030383; BAC26935.1; -.
DR MEROPS; C19.075; -.
DR MGD; MGI:1859645; Usp27X.
DR GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004221; F:ubiquitin thiolesterase activity; IEA.
DR GO; GO:0006511; P:ubiquitin-dependent protein catabolism; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001394; Peptidase_C19.
DR InterPro; IPR001607; Znf_UBP.
DR Pfam; PF00443; UCH; 1.
DR Pfam; PF02148; zf-UBP; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE; PS00972; UCH_2_1; 1.
DR PROSITE; PS00973; UCH_2_2; 1.
DR PROSITE; PS02335; UCH_2_3; 1.
DR Protease.
KW NON_TER.
FT TER.
SQ SEQUENCE 540 AA; 61206 MW; E0F4C3E840A0A4A CRC64;

Query Match 90.6%; Score 29; DB 2; Length 540;
Best Local Similarity 80.0%; Pred. No. 9.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
DB 282 CKGDD 286

RESULT 73
Q98UH5 PRELIMINARY; PRT; 549 AA.
AC Q98UH5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)

```

```

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
GN Name=GRK7;
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxID=7962;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RA Shimauchi Y., Tachibanaki S., Kawamura S.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. GPRK
CC subfamily.
CC -!- SIMILARITY: Contains 1 RGS domain.
DR EMBL; AB055658; BAB32498.1; -.
DR HSSP; P31751; IGZK.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004703; F:G-protein coupled receptor kinase activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0006740; F:transferase activity; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR000239; GPCR_Kinase.
DR InterPro; IPR011009; Kinase_like.
DR InterPro; IPR000961; Pkinase_C.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR000342; RGS.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00069; Pkinase; 1.
DR Pfam; PF00615; RGS; 1.
DR PRINTS; PR007117; GPCR_KINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00315; RGS; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00133; S_TK_X; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00132; RGS; 1.
KW ATP-binding; Kinase; Receptor; Serine/threonine-protein kinase;
KW Transferase.
SQ SEQUENCE 549 AA; 62423 MW; CB186B8587A0B183 CRC64;

Query Match 90.6%; Score 29; DB 2; Length 549;
Best Local Similarity 80.0%; Pred. No. 9.4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
DB 438 CKGDD 442

RESULT 74
Q99KU3 PRELIMINARY; PRT; 553 AA.
AC Q99KU3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein OSUNBA0006G10.30 (Hypothetical protein
DE P0584D02.22).
GN Name=OSUNBA0006G10.30; Synonyms=P0584D02.22;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]

```

RP SEQUENCE FROM N.A.
 RA Sasaki T., Matsumoto T., Katayose Y.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 9, BAC
 RT clone:OSJNBa006G10."
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Sasaki T., Matsumoto T., Katayose Y.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 9, PAC
 RT clone:P0584D02."
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP005909; BAD34163.1; -
 DR EMBL; AP005634; BAD33695.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 553 AA; 60475 MW; 24E5AF1DADA67D4 CRC64;

Query Match 90.6%; Score 29; DB 2; Length 553;
 Best Local Similarity 80.0%; Pred.No. 9.5e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
 |::||
 Db 289 CKGDD 293

RESULT 75

Q6DCJ1 PRELIMINARY; PRT; 572 AA.
 AC Q6DCJ1;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Hypothetical protein.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
 OC Xenopodinae; Xenopus.
 ON NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.B., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative."
 RL Dev. Dyn. 225:384-391(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.T., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Boeak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski N.I., Skalska U., Smallos D.E., Schmerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP SEQUENCE FROM N.A.

RC TISSUE=Embryo;
 RA Klein S., Strausberg R.;
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC078033; AAH78033.1; -
 DR GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.
 DR GO; GO:0004221; F:ubiquitin thiolesterase activity; IEA.
 DR GO; GO:0006511; F:ubiquitin-dependent protein catabolism; IEA.
 DR InterPro; IPR000345; CytC_heme_BS.
 DR InterPro; IPR001394; Peptidase_C19.
 DR Pfam; PF00443; UCH; 1.
 DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_2.
 DR PROSITE; PS00972; UCH_2_1; 1.
 DR PROSITE; PS00973; UCH_2_2; 1.
 DR PROSITE; PS00235; UCH_2_3; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 572 AA; 65376 MW; 2877008E74EB7B6B CRC64;

Query Match 90.6%; Score 29; DB 2; Length 572;
 Best Local Similarity 80.0%; Pred.No. 9.8e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CRGDD 5
 |::||
 Db 325 CKGDD 329

Search completed: September 7, 2005, 20:06:55
 Job time : 202 secs

This Page Blank (uspto)